

Disentangling Clustering Effects in Jet Algorithms

Randall Kelley

Department of Physics, Harvard University, Cambridge, Massachusetts, 02138, USA
E-mail: rkelley@physics.harvard.edu

Jonathan R. Walsh and Saba Zuberi

*Theoretical Physics Group, Ernest Orlando Lawrence Berkeley National Laboratory,
and Center for Theoretical Physics, University of California, Berkeley, CA 94720, USA*
E-mail: jwalsh@lbl.gov, szuberi@lbl.gov

ABSTRACT: Clustering algorithms build jets through the iterative application of single particle and pairwise metrics. This leads to phase space constraints that are extremely complicated beyond the lowest orders in perturbation theory, and in practice they must be implemented numerically. This complication presents a significant barrier to gaining an analytic understanding of the perturbative structure of jet cross sections. We present a novel framework to express the jet algorithm's phase space constraints as a function of clustered groups of particles, which are the possible outcomes of the algorithm. This approach highlights the analytic properties of jet observables, rather than the explicit constraints on individual final state momenta, which can be unwieldy at higher orders. We derive the form of the n -particle phase space constraints for a jet algorithm with any measurement. We provide an expression for the measurement that makes clustering effects manifest and relates them to constraints from clustering at lower orders. The utility of this framework is demonstrated by using it to understand clustering effects for a large class of jet shape observables in the soft/collinear limit. We apply this framework to isolate divergences and analyze the logarithmic structure of the Abelian terms in the soft function, providing the all-orders form of these terms and showing that corrections from clustering start at next-to-leading logarithmic order in the exponent of the cross section.

KEYWORDS: [Jets](#), [Jet Algorithms](#).

Contents

1. Introduction	1
1.1 The Complication of Measuring Properties of Jets	2
1.2 Reformulating the Constraints from Clustering	4
1.3 An Application: Soft Clustering and Jet Shapes	5
1.4 Outline	8
2. Phase Space Constraints from Clustering Algorithms	9
2.1 Unitarity Relations and Phase Space Constraints	9
2.2 Clustering Maps	12
3. The Measurement Function for Clustering Algorithms	16
3.1 Main Results	21
4. Cancellation of Infrared Divergences from Clustering	22
5. An Application: All-Orders Abelian Structure of Soft Clustering for Jet Shapes	23
5.1 All-Orders Abelian Structure	25
5.2 Renormalization of the Abelian Soft Function	27
5.3 Failure of Exponentiation	28
6. Conclusions	29
7. Acknowledgments	29
A. Examples for 4 and 5 Final State Particles	30

1. Introduction

Jet algorithms play an essential role in high energy collision experiments, organizing the hadronic structure of an event to allow for a meaningful interpretation of the short distance interaction. The LHC has increased interest in understanding and utilizing the structure of jets [1, 2], and these tools often make use of clustering algorithms such as anti- k_T , Cambridge/Aachen (C/A), and k_T [3, 4, 5, 6, 7]. Understanding the structure of jets requires understanding their perturbative description.

Clustering algorithms build jets through an iterative recombination procedure that sequentially identifies the pair of particles that are closest according to a distance measure and then *clusters* them, i.e. combines their momenta, doing so until a stopping criterion

is reached. Each algorithm differs mainly in the choices of the distance measure and the stopping criterion used in this procedure. Implementing phase space constraints from clustering for final states with more than two or three particles is very complicated, and can present an impasse to analytic perturbative calculations, which in many cases must be performed numerically. This makes it difficult to understand the effects of clustering on the perturbative series in jet cross sections.

1.1 The Complication of Measuring Properties of Jets

When measuring a set of observables, $\{O\}$, clustering can strongly affect their values by pulling particles into and out of the various regions where $\{O\}$ is measured. When calculating these effects on the differential cross section, it is useful to separate the calculation into two parts – the squared matrix element, \mathcal{A} , and the measurement function, \mathcal{M} , which implements the measurement of $\{O\}$ on the final state. The measurement function depends on the final state momenta at a given point in phase space, Φ , and the observables being measured. It is integrated against \mathcal{A} when evaluating the contribution to the cross section:

$$\frac{d\sigma}{d\{O\}} = \int d\Phi \mathcal{A}(\Phi) \mathcal{M}(\{O\}, \Phi). \quad (1.1)$$

We will be concerned with understanding the general structure of the measurement function for a wide range of observables for jets defined by a clustering algorithm. This requires translating the iterative recombination procedure of clustering algorithms into all-orders properties of the measurement function – it is a nontrivial task.

The measurement function is composed of two main parts: a function of momenta, $f_r(\{k\})$, associated with the measured observable $O_r \in \{O\}$ and a restriction, $\mathcal{R}_r(\{k\}, \Phi)$, on the region of phase space over which we measure it. Here $\{k\}$ is a subset of all momenta for a given phase space point, Φ , that \mathcal{R}_r requires to be in region r . The measurement function can be expressed schematically as

$$\mathcal{M}(\{O\}, \Phi) = \prod_r \delta\left(O_r - \sum_{\{k\}} f_r(\{k\}) \mathcal{R}_r(\{k\}; \Phi)\right), \quad (1.2)$$

where the product is over the regions of phase space, r , which contribute to the cross section. For example, the constraints that contribute to a dijet cross section are illustrated in Fig. 1 and correspond to the two jets, \mathcal{R}_{J_1} and \mathcal{R}_{J_2} , and an out-of-jet region, \mathcal{R}_{out} , where a different function of momenta, $f_r(\{k\})$, may be measured in each region. It is often simpler to work with the Fourier transform of the measurement function with respect to the observables:

$$\begin{aligned} \widetilde{\mathcal{M}}(\{x\}, \Phi) &= \prod_r \int_{-\infty}^{\infty} dO_r e^{-ix_r O_r} \mathcal{M}(\{O\}, \Phi) \\ &= \prod_r \exp\left(-ix_r \sum_{\{k\}} f_r(\{k\}) \mathcal{R}_r(\{k\}; \Phi)\right). \end{aligned} \quad (1.3)$$

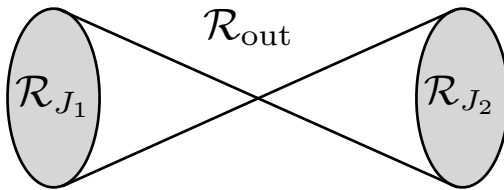


Figure 1: Phase space regions \mathcal{R}_r over which observables O_r are measured, contributing to a dijet measurement function.

An illustrative example is the measurement of the invariant mass of a jet, m_J^2 ; it contributes a factor to \mathcal{M} of

$$\mathcal{M}(m_J^2, \Phi) = \delta \left(m_J^2 - \sum_{i,j} k_i \cdot k_j \mathcal{R}_J(\{k_i, k_j\}; \Phi) \right), \quad (1.4)$$

where $\mathcal{R}_J(\{k_i, k_j\}; \Phi)$ gives the phase space constraints for particles i, j to be in the jet J . Expressed explicitly, \mathcal{R}_J is given by a product of theta functions of momenta in Φ that constrain k_i and k_j to be in the jet. In general, whether k_i and k_j are in the jet depends on the clustering procedure of the jet algorithm, meaning \mathcal{R}_J depends on all final state momenta Φ . While the form of Eq. (1.2) is compact, it hides correlations between particles, since \mathcal{R}_J depends on momenta not in the jet. When the measurement function is used in fixed order calculations, the hidden phase space constraints from clustering must be made more explicit. However, the form of Eq. (1.2) makes the perturbative structure of the jet cross section difficult to understand analytically without explicitly carrying out higher order calculations.

The JADE algorithm [8] provides an example of why it is important to make manifest the effects of clustering that are hidden in \mathcal{R}_r in Eq. (1.2). Although JADE is a well-defined infrared (IR) safe algorithm that combines particles based on an invariant mass metric, clustering effects at higher orders change the structure of the leading log series ($\alpha_s^n \ln^{2n}$) and spoil resummation at this order. This fact is obscured by expressing the action of the algorithm solely through the iterative application of its metric. It required an analytic calculation at two loops [9] to understand the perturbative structure of jet cross sections defined with JADE. More recently it has been shown that clustering effects for the C/A and k_T algorithms change the structure of the perturbative series at $\mathcal{O}(\alpha_s^2 \ln^2)$ by explicit calculation of the Abelian terms, likely spoiling resummation at this order. Such effects are absent for anti- k_T , providing a theoretical preference for this algorithm.

These examples illustrate the importance of having an analytic understanding of the perturbative structure of jet cross sections. However higher order analytic calculations are difficult and can be made prohibitively complex by clustering phase space constraints. Since the structure of the measurement function, \mathcal{M} , determines clustering effects, we focus on expressing \mathcal{M} in a form that makes these effects manifest. This will provide

insight into the perturbative structure of jet cross sections without having to carry out explicit calculations.

1.2 Reformulating the Constraints from Clustering

Running a clustering jet algorithm divides up phase space into mutually exclusive regions that cover all of phase space. For an N -jet observable these regions may correspond to each jet, $\mathcal{R}_{J_1}, \dots, \mathcal{R}_{J_N}$, and an out-of-jet region \mathcal{R}_{out} . The jet algorithm does this by separating final state particles into groups that characterize the outcomes of the algorithm. For example, for three particles with momentum k_1 , k_2 , and k_3 , the possible outcomes, $\mathcal{K}_\ell(\Phi)$, of an algorithm are

- k_1 , k_2 , and k_3 all cluster together: \mathcal{K}_3
- k_1 and k_2 cluster, but k_3 does not (+ permutations): $\mathcal{K}_{12}, \mathcal{K}_{13}$, and \mathcal{K}_{23}
- None of the particles cluster: \mathcal{K}_1

Each outcome $\mathcal{K}_\ell(\Phi)$ represents a set of phase space constraints requiring the particles to cluster into a set of groups. Summing over all possible outcomes, ℓ , of a clustering algorithm at a given number of final state particles gives a unitarity relation:

$$1 = \sum_{\ell} \mathcal{K}_\ell(\Phi). \quad (1.5)$$

Since the \mathcal{K}_ℓ makes the phase space constraints for a given outcome of the clustering algorithm explicit, \mathcal{R}_r is simplified. $\mathcal{R}_r(\{k\}; \Phi)$ constrains $\{k\}$ to be in region r where the observable O_r is measured, but it no longer must implement the constraints from clustering, which appear in the \mathcal{K}_ℓ . This means \mathcal{R}_r depends only on the momenta in the region r :

$$\mathcal{R}_r(\{k\}; \Phi) \rightarrow \mathcal{R}_r(\{k\}). \quad (1.6)$$

To write the measurement function, we simply weight Eq. (1.5) by Eq. (1.2):

$$\mathcal{M}(\{O\}, \Phi) = \sum_{\ell} \mathcal{K}_\ell(\Phi) \prod_r \delta\left(O_r - \sum_{\{k\}} f_r(\{k\}) \mathcal{R}_r(\{k\})\right). \quad (1.7)$$

While it is simple to enumerate the outcomes of clustering, each \mathcal{K}_ℓ contains a complicated set of phase space constraints; as such, in order to explore this structure we need to know where the IR divergences arise, since they can contribute large logarithms in the perturbative series of the cross section. This depends on both the squared matrix element, \mathcal{A} , and the measurement, \mathcal{M} . Although the structure of \mathcal{A} is complicated at higher orders, we will see that examining the structure of \mathcal{M} alone can provide insight. The differential cross section is sensitive to the IR regions of phase space determined by \mathcal{R}_r in Eq. (1.3).

In this paper we provide a framework to use unitarity relations of the type Eq. (1.5) to express the outcome of the clustering algorithm and isolate the most divergent regions of phase space, which contribute the most logarithms in the perturbative expansion. As we shall see, this is done by using unitarity iteratively to rewrite *exclusive* constraints as *inclusive* ones. For example, the case where k_1 and k_2 cluster but k_3 does not contains a

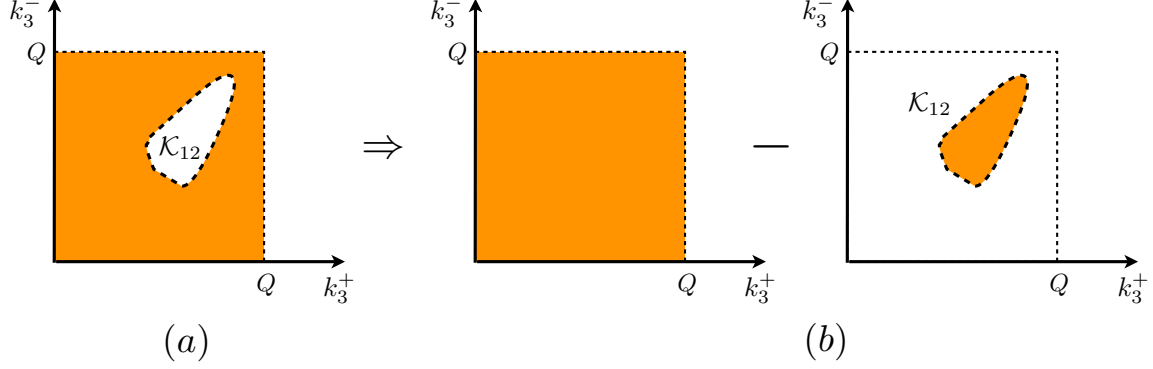


Figure 2: Phase space for particle k_3 for fixed k_1, k_2 . In the region \mathcal{K}_{12} , k_3 would be clustered with k_1 and k_2 . The exclusive constraint $\theta(k_3 \notin \mathcal{K}_{12})$ is shown in (a) by the orange region and the inclusive constraint $1 - \theta(k_3 \in \mathcal{K}_{12})$ is shown in (b) by the orange region.

phase space constraint excluding k_3 from some region of phase space; it cannot be close enough to k_1 or k_2 to fall in one of the other outcomes. This is shown in Fig. 2(a), where for a given k_1 and k_2 there is a region \mathcal{K}_{12} that k_3 can not occupy in order to avoid being clustered¹. Divergences in phase space arise as $k_3^+, k_3^- \rightarrow 0$ along the axes and the associated logarithms will depend on k_1 and k_2 in a nontrivial way through the boundary of \mathcal{K}_{12} . A more useful description of clustering would rewrite this in terms of an inclusive phase space constraint shown in Fig. 2(b), where the exclusive and inclusive ways to write this constraint would be

$$\begin{aligned} \text{exclusive : } & \theta(k_3 \notin \mathcal{K}_{12}), \\ \text{inclusive : } & 1 - \theta(k_3 \in \mathcal{K}_{12}). \end{aligned} \tag{1.8}$$

The function $\theta(x)$ is the usual step function, giving 1 if $x > 0$ and 0 otherwise. The inclusive description expresses the divergences in k_3 in a way that is independent of k_1 and k_2 and subtracts a finite contribution from the region \mathcal{K}_{12} . As a result, integrating over the momenta of particle k_3 with the inclusive constraints simplifies the divergent structure [10, 11]. Rewriting the exclusive constraint as an inclusive one uses a (simple) unitarity relation describing the phase space constraints between particles. One can iteratively apply unitarity relations of this kind to rewrite the outcome of clustering only in terms of inclusive phase space constraints.

1.3 An Application: Soft Clustering and Jet Shapes

Energetic particles have the ability to change the number, energy, and direction of jets. Often, though, energetic particles are confined to small regions of phase space and only

¹We decompose the four-momentum as

$$k^\mu = n \cdot k \frac{\bar{n}^\mu}{2} + \bar{n} \cdot k \frac{n^\mu}{2} + k_\perp^\mu,$$

and define $k^+ \equiv n \cdot k$ and $k^- \equiv \bar{n} \cdot k$. Here n and \bar{n} are light-cone vectors, $n^\mu = (1, \hat{n})$ and $\bar{n}^\mu = (1, -\hat{n})$, with $\hat{n}^2 = 1$.

cluster within jets. This occurs for jet observables where the measurement of an observable O constrains the collinear radiation to be tightly collimated within its own jet, compared to the jet size; e.g. measuring the jet mass, m_J , when $m_J/E_J \ll R$ forces this to be true. This is satisfied for a range of jet shapes and substructure observables [12, 13, 14, 15, 16] in the soft/collinear regime. In this case the boundaries of the jets are determined by clustering amongst *soft* particles.

In this regime the function of momenta that we measure, $f_r(\{k\})$ in Eq. (1.3), simplifies to be a function of momentum of a *single* particle $f_r(k_i)$ for a large class of observables. For the example of jet mass given in Eq. (1.4), the expression greatly simplifies to give

$$\sum_{i,j} k_i \cdot k_j \mathcal{R}_J(\{k_i, k_j\}) \rightarrow \sum_i p_J \cdot k_i \mathcal{R}_J(k_i), \quad (1.9)$$

where p_J is the sum of the collinear momentum and is conserved for each jet direction in the soft/collinear regime. The measurement function still contains correlations between particles through the outcomes of the algorithm, given by \mathcal{K}_ℓ :

$$\widetilde{\mathcal{M}}(\{x\}, \Phi) = \sum_\ell \mathcal{K}_\ell(\Phi) \prod_r \exp \left(-i \sum_i x_r f_r(k_i) \mathcal{R}_r(k_i) \right). \quad (1.10)$$

Soft clustering for the anti- k_T algorithm is particularly simple: jets in the soft/collinear regime are circular with jet size R around the collinear direction up to power suppressed corrections [7, 13, 17]. In this case the sum over outcomes in Eq. (1.10) simplifies to a single outcome: none of the particles cluster, and the constraint is trivial, $\mathcal{K}(\Phi) = 1$. This removes all correlations between different particles in the final state from the measurement function. As a result the measurement simplifies greatly:

$$\widetilde{\mathcal{M}}_{\text{akT}}(\{x\}, \Phi) = \prod_i \widetilde{\mathcal{M}}^{(1)}(\{x\}, k_i). \quad (1.11)$$

where

$$\widetilde{\mathcal{M}}^{(1)}(\{x\}, k) = \exp \left(-i \sum_r x_r f_r(k) \mathcal{R}_r(k) \right). \quad (1.12)$$

It is well known that the measurement function also takes this simple form for a large class of event shapes in the soft/collinear limit, such as thrust, heavy jet mass, angularities, and N-jettiness [18, 19, 20, 21], which makes such shapes theoretically appealing. Notice that jet rates do not satisfy the form in Eq. (1.11) except for fixed cone algorithms, which have known split/merge problems [22, 23].

Algorithms such as C/A or k_T modify the phase space regions from the simple cones of the anti- k_T algorithm, due to clustering amongst soft gluons. This occurs when a soft particle is pulled in or out of the jet, and we refer to this as

boundary clustering : clustering of soft particles across the jet boundary of radius R
around the jet direction.

Example configurations for 2 and 3 particles are shown in Fig. 3. Since boundary clustering

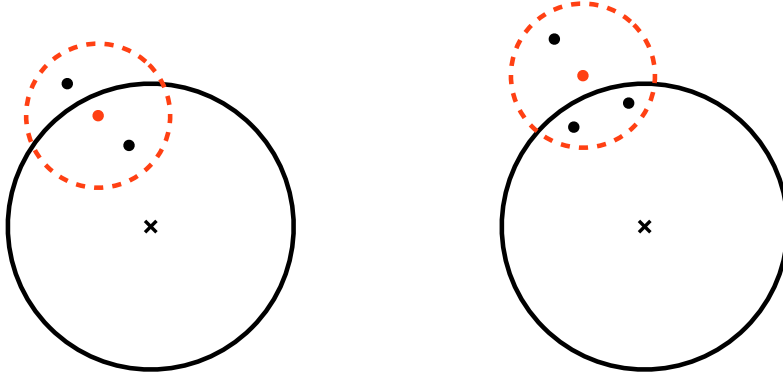


Figure 3: Configurations of two (left) and three (right) particle boundary clustering that changes the jet boundary. The final state particles shown in black cluster, and the red particle is the result. In the two particle case the particle outside the jet is pulled in by clustering, and in the three particle case the particles in the jet are pulled out.

modifies the region over which the measurement is made and is an effect that intrinsically depends on the momenta of all particles that are boundary clustered, the phase space constraints for the outcome, $\mathcal{K}_\ell(\Phi)$, of the C/A or k_T measurement functions will depend on multiple momenta in the final state. This introduces correlations into the measurement function, which in this case is of the form Eq. (1.10). A useful way to organize the correlations introduced by $\mathcal{K}_\ell(\Phi)$ is through a correction to the measurement function of the form

$$\widetilde{\mathcal{M}}^{(n)}(\{x\}, \Phi) = \prod_{i=1}^n \widetilde{\mathcal{M}}^{(1)}(\{x\}, k_i) + \widetilde{\mathcal{M}}_{\text{corr.}}^{(n)}(\{x\}, \Phi), \quad (1.13)$$

where $\widetilde{\mathcal{M}}^{(1)}$ is given in Eq. (1.12) and n is the number of final state particles. This is obtained from Eq. (1.10) by a trivial application of the unitarity relation $\mathcal{K}_\ell(\Phi) \equiv 1 - \overline{\mathcal{K}}_\ell(\Phi)$ for each outcome ℓ , where this relation defines $\overline{\mathcal{K}}_\ell(\Phi)$.

Clustering for algorithms other than anti- k_T affect the logarithmic terms arising from soft divergences. In the soft and collinear regimes, boundary clustering can lead to large logarithms in the cross section, starting at $\mathcal{O}(\alpha_s^2)$. This was shown in [24, 25] and further explored in [26, 27], and the original authors termed the effect *clustering logs*. In [27] it was shown that clustering logs have the same properties as non-global logarithms (NGLs) [28, 29] in the soft/collinear limit and that they contribute at least at next-to-leading logarithm² (NLL), but could in principle contribute at LL. It was also argued that new clustering effects likely arise at each order in α_s at NLL and beyond, spoiling resummation.

Using unitarity relations of the form Eq. (1.5), we can express the measurement function to isolate the divergences from boundary clustering. These corrections can be parameterized in terms of the number of particles involved in clustering. In particular we

²We count in the exponent of the distribution, where N^kLL terms are of order $\alpha_s^n \ln^{n-k+1}$.

show

$$\widetilde{\mathcal{M}}_{\text{corr.}}^{(n)}(\{x\}, \Phi) = \sum_{k=0}^{n-2} \frac{n!}{(n-k)! k!} \Delta \widetilde{\mathcal{M}}^{(n-k)} [\widetilde{\mathcal{M}}^{(1)}]^k, \quad (1.14)$$

where $\Delta \widetilde{\mathcal{M}}^{(k)}$ is the contribution where k particles cluster with at least one other particle and first arises at $\mathcal{O}(\alpha_s^k)$. This result applies not only to boundary clustering, but clustering in general, away from soft-collinear regime. While clustering effects can a priori contribute at LL, the form of the measurement function in Eq. (1.14) allows us to show that clustering logarithms arise at NLL and to determine to all orders the form of the Abelian soft function and show that exponentiation of these terms does not occur.

1.4 Outline

In this paper, we construct a framework to describe the phase space constraints from clustering and the effect on observables. The formalism we use is very general, applying to a wide class of clustering algorithms and measurements. It can be applied to a specific type of clustering, such as boundary clustering, or to clustering in general. The formalism to describe clustering is presented in two main parts.

First, in Sec. 2, we define a set of functions that schematically describe the phase space constraints from clustering. They are functions of clustered groups of particles, which are the possible outcomes of the jet algorithm. The physical interpretation of these functions allows us to construct unitarity relations that transform exclusive constraints into inclusive ones. The result is a prescription to systematically determine the phase space constraints from clustering in terms of inclusive constraints. Our goal is not to write clustering constraints explicitly in terms of the momenta of individual particles, but instead to express the constraints in a form that highlights the analytic properties of the outcome of a generic clustering algorithm. We do not deal directly with the individual clustering steps of the algorithm, but in most cases our results can be applied to describe each step.

Second, in Sec. 3 we apply this framework to the measurement function. For any measurement function of the form Eq. (1.7) where the observables depend on the momenta of a single clustered group, we use the results of Sec. 2 to find the corrections due to clustering. We show that these have the form Eq. (1.14). This form can be exploited to determine the organization of logarithms arising from clustering. We also comment on the infrared safety of clustering algorithms in Sec. 4; we can use this framework to see precisely how clustering effects among soft particles are IR safe.

In Sec. 5, we apply the formalism of the measurement function to study the logarithmic structure of the soft function arising from clustering. We focus on the Abelian terms, where the matrix elements factorize and it is straightforward to translate the measurement function into contributions to the cross section. We use the form of the measurement function along with constraints from renormalization to determine the all-orders structure of the soft function, allowing us to show that clustering NGLs contribute at most at NLL at all orders in α_s . Finally, in Sec. 6, we summarize our results.

2. Phase Space Constraints from Clustering Algorithms

In this section and the next, “clustering” can refer to general clustering or to a specific kind, such as boundary clustering. If this framework is used to describe boundary clustering, which is relevant for observables that constrain the collinear radiation in the jet to be collimated compared to the jet size (e.g. $m_J/E_J \ll R$), then the phase space constraints will ignore clustering that does not change the jet boundary. We will also treat the particles as indistinguishable; this is not necessary, but simplifies the discussion.

As discussed in the introduction, the deterministic nature of jet algorithms allows the action of the jet algorithm to be characterized in terms of a set of mutually exclusive outcomes, \mathcal{K}_ℓ , as in Eq. (1.5). When we run the jet algorithm, sets of particles will cluster, and we use these sets to characterize the outcomes of the algorithm. For example, if there are 5 particles in the final state labeled q_1, \dots, q_5 , then a possible outcome from a jet algorithm \mathcal{J} is

$$\mathcal{J} : \{q_1, q_2, q_3, q_4, q_5\} \longrightarrow \{\{q_1, q_2\}, \{q_3, q_4\}, \{q_5\}\}. \quad (2.1)$$

In this case particles 1 and 2 have been clustered, as have particles 3 and 4, and particle 5 does not cluster.

This example shows that the action of a jet algorithm is to partition the final state particles into groups. We refer to partitions like the one above as labeled partitions. There is a one-to-one association between labeled partitions and outcomes of the algorithm. However, if the particles are indistinguishable, there are many labeled partitions that describe the same basic outcome. In the case of Eq. (2.1), the important characteristic is that particles were clustered in groups of 2, 2, and 1; a distinct clustering that is equivalent to this outcome is $\{\{q_1, q_4\}, \{q_2, q_5\}, \{q_3\}\}$. Therefore, we can simplify the labeled partitions and describe them in terms of the number of particles clustered in each group. We will refer to these simply as partitions. The labeled partition in Eq. (2.1) maps to the partition $\{2, 2, 1\}$. For 5 particles, the possible partitions are

$$\{5\}, \{4, 1\}, \{3, 2\}, \{3, 1, 1\}, \{2, 2, 1\}, \{2, 1, 1, 1\}, \{1, 1, 1, 1, 1\}. \quad (2.2)$$

In general, we will denote partitions by $\mathcal{P} = \{p_1, \dots, p_k\}$, where the p_i are natural numbers. The number of labeled partitions that are equivalent to the partition \mathcal{P} is

$$N(\mathcal{P}) = \frac{n!}{p_1! \dots p_k!} \frac{1}{d_1! \dots d_n!}, \quad (2.3)$$

where n is the total number of particles, and d_i is the number of elements of \mathcal{P} equal to i . The first factor in $N(\mathcal{P})$ is the usual multinomial coefficient, and the second factor accounts for the degeneracy in dividing particles among equal sized elements in the partition.

2.1 Unitarity Relations and Phase Space Constraints

We will now discuss how we can more specifically determine the phase space constraints \mathcal{K}_ℓ , defined in Sec. 1, for each outcome, ℓ , of the jet algorithm (which we associate with a

partition). In words, a partition represents the phase space constraints:

$$\{p_1, \dots, p_k\} : p_1 \text{ particles cluster, } p_2 \text{ particles cluster, } \dots, p_k \text{ particles cluster,} \\ \text{and these groups of particles do not cluster with each other.} \quad (2.4)$$

Therefore, we must define two basic functions that act on the elements of a partition: one requiring a set of particles to cluster and one requiring different groups of particles not to cluster. These can be defined in terms of objects in the partition:

$$c(p) : p \text{ particles are clustered,} \quad (2.5)$$

and

$$s(\{p_1, \dots, p_k\}) : \text{the clusters of } p_1, \dots, p_k \text{ particles do not cluster.} \quad (2.6)$$

For completeness, we define $s(\{p\}) \equiv 1$ and $c(1) \equiv 1$ (i.e., no constraint). The function $s(\{p_1, \dots, p_k\})$ defines an *exclusive* constraint; for example, it excludes the set of particles p_k from some region of phase space so that they do not combine with the sets of particles p_1, \dots, p_{k-1} . For a given partition \mathcal{P} , the phase space constraints we assign to it are

$$\left[\prod_{p \in \mathcal{P}} c(p) \right] s(\mathcal{P}). \quad (2.7)$$

If we sum over all possible outcomes of the jet algorithm, we obtain a unitarity relation for phase space:

$$1 = \sum_{\mathcal{P}} N(\mathcal{P}) \left[\prod_{p \in \mathcal{P}} c(p) \right] s(\mathcal{P}), \quad (2.8)$$

where the sum is over partitions. Each term in the sum corresponds to an outcome \mathcal{K}_ℓ of the algorithm, so that Eq. (2.8) is a more specific version of Eq. (1.5).

Some basic examples are useful. For two particles, the partitions are $\{1, 1\}$ and $\{2\}$, and the unitarity relation is

$$1 = c(2) + s(\{1, 1\}). \quad (2.9)$$

This relation defines $s(\{1, 1\})$. For three particles, the partitions are $\{1, 1, 1\}$, $\{1, 2\}$, and $\{3\}$, and the unitarity relation is

$$1 = c(3) + 3 c(2) s(\{1, 2\}) + s(\{1, 1, 1\}). \quad (2.10)$$

Consider the function $s(\{1, 2\})$. Let us label the single particle q_1 and the pair of particles that cluster q_2 and q_3 . The function $s(\{1, 2\})$ allows q_1 to go anywhere except for near q_2 and q_3 . However, we can use a unitarity constraint to rewrite $s(\{1, 2\})$ in terms of an *inclusive* constraint:

$$s(\{1, 2\}) = (q_1 \text{ goes anywhere}) - (q_1 \text{ would cluster with } q_2, q_3, \text{ or } q_2 + q_3). \\ \equiv 1 - m(\{1, 2\}). \quad (2.11)$$

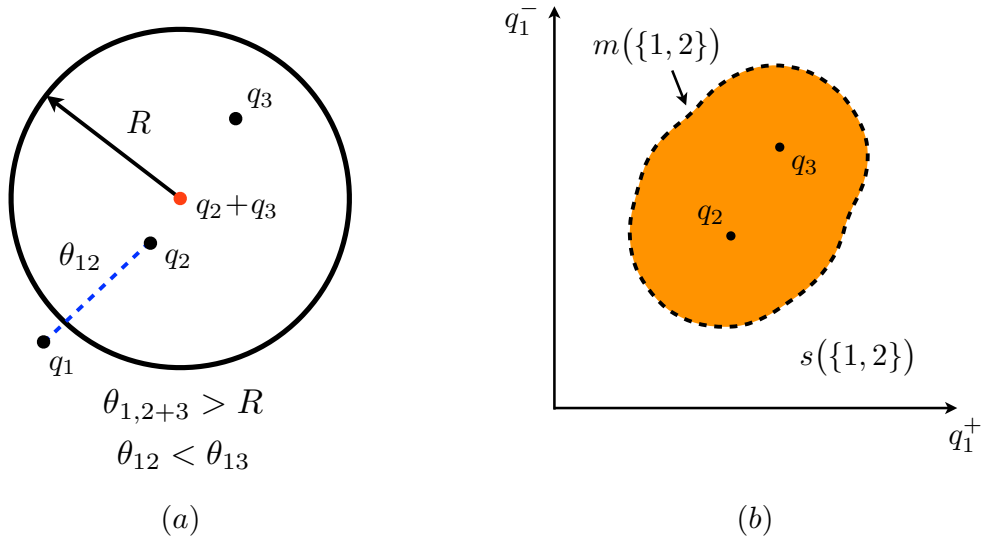


Figure 4: On the left, a configuration of 3 particles for the C/A algorithm that contributes to $c(2)m(\{1,2\})$ but not to $c(3)$. In this case particles q_2 and q_3 cluster first to form the merged pair $q_2 + q_3$, shown in red. Particle q_1 is too far away to cluster with $q_2 + q_3$, as it lies outside the solid circle of radius R centered at $q_2 + q_3$. Therefore the 3 particles do not cluster, and so do not contribute to $c(3)$. On the right, a schematic picture of the region of phase space for particle q_1 defined by $m(\{1,2\})$, which is the complement of $s(\{1,2\})$.

This relation defines the *merging* function $m(\{1,2\})$, which requires q_1 to be near the clustered pair. However, this function does not require that all three particles would cluster when running the jet algorithm:

$$c(2)m(\{1,2\}) \neq c(3). \quad (2.12)$$

For example, for fixed momenta of q_2 and q_3 , the single particle q_1 could be close enough to q_2 but sufficiently far from q_3 so that all three particles would not cluster. In Fig. 4, we show a configuration that is part of $c(2)m(\{1,2\})$ but not part of $c(3)$, as well as a schematic region of phase space for $m(\{1,2\})$. Only in the case of two particles does $m(\{1,1\}) = c(2)$. In general, merging requires that at least one particle from each group would cluster with at least one particle in another group, such that all the groups are connected. The general merging function is

$$m(\{p_1, \dots, p_k\}) : \text{the clusters of } p_1, \dots, p_k \text{ particles merge.} \quad (2.13)$$

Our general goal will be to rewrite the exclusive $s(\mathcal{P})$ functions in terms of the inclusive clustering and merging functions that require particles to be grouped together. For each function $s(\mathcal{P})$, we can define a unitarity relation to rewrite $s(\mathcal{P})$ in terms of clustering and merging functions. Note that rewriting $s(\mathcal{P})$ does not change the outcome of the jet algorithm, but instead rewrites the phase space constraints from clustering in the measurement function in a way that makes the perturbative structure of the cross section clearer.

We introduce some diagrammatic notation which allows us to express these unitarity relations in a simple way. These diagrams are helpful in thinking about higher order particle configurations and are useful in determining unitarity relations. Instead of clustering in 3-space, it is sufficient to think of clustering along a line. Each particle will be depicted as a dot, and if particles are clustered then we circle them. If the particles do not cluster, we put a vertical bar between them. If particles are not on the same line, then their phase space constraints do not depend on each other. Finally, when we require groups of particles to merge, we circle those groups and place a vertical bar between the groups. In Fig. 5, we give the diagrams for the 2 and 3 particle cases.

$$\begin{aligned}
& \mathbf{n = 2} \\
& \begin{array}{c} \bullet \\ \hline \bullet \\ \hline \end{array} \quad \underset{1}{=} \quad \begin{array}{c} \bullet \bullet \\ \hline \end{array} \quad \underset{c(2)}{=} \quad + \quad \begin{array}{c} \bullet \quad | \quad \bullet \\ \hline \end{array} \quad \underset{s(\{1,1\})}{=}
\end{aligned}$$

$$\begin{aligned}
& \mathbf{n = 3} \\
& \begin{array}{c} \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array} \quad \underset{1}{=} \quad \begin{array}{c} \bullet \bullet \bullet \\ \hline \end{array} \quad \underset{c(3)}{=} \quad + \quad \begin{array}{c} \bullet \bullet \bullet \quad | \quad \bullet \\ \hline \end{array} \quad \underset{3\,c(2)s(\{1,2\})}{=} \quad + \quad \begin{array}{c} \bullet \quad | \quad \bullet \quad | \quad \bullet \\ \hline \end{array} \quad \underset{s(\{1,1,1\})}{=}
\end{aligned}$$

$$\begin{aligned}
& \text{where} \quad \begin{array}{c} \bullet \bullet \bullet \quad | \quad \bullet \\ \hline \end{array} \quad \underset{c(2)s(\{1,2\})}{=} \quad = \quad \begin{array}{c} \bullet \bullet \bullet \\ \hline \bullet \end{array} \quad \underset{c(2)}{=} \quad - \quad \begin{array}{c} \bullet \bullet \bullet \quad | \quad \bullet \bullet \bullet \\ \hline \end{array} \quad \underset{c(2)m(\{1,2\})}{=}
\end{aligned}$$

Figure 5: Clustering diagrams for the cases $n = 2$ and $n = 3$. Particles are represented by dots, and clustered particles are circled. Particles that do not cluster have a vertical bar between them. Merged groups of clustered particles are circled, with a vertical bar between each group. The different terms in the unitarity relations are shown. We also show how to express the $s(\mathcal{P})$ functions requiring particles to not cluster in terms of $c(p)$ and $m(\mathcal{P})$ functions that require the particles to be clustered/merged.

2.2 Clustering Maps

When we rewrite the phase space constraints given by $s(\mathcal{P})$ in Eq. (2.8), we can represent the result as a map between partitions. Consider the unitarity relation that defines $s(\{1, 1, 1\})$ in Fig. 5:

$$\begin{aligned}
s(\{1, 1, 1\}) &= 1 - c(3) - 3\,c(2)s(\{1, 2\}) \\
&= 1 - c(3) - 3\,c(2) + 3\,c(2)m(\{1, 2\}).
\end{aligned} \tag{2.14}$$

Each term represents a different phase space constraint on the three particles, and correspondingly a map to a different partition.

$$\begin{array}{rclcl}
1 & : & \{1, 1, 1\} & \rightarrow & \{1, 1, 1\} \\
-3c(2) & : & \{1, 1, 1\} & \rightarrow & \{1, 2\} \\
3c(2)m(\{1, 2\}) & : & \{1, 1, 1\} & \rightarrow & \{1, 2\} \rightarrow \{3\} \\
-c(3) & : & \{1, 1, 1\} & \rightarrow & \{3\}
\end{array} \tag{2.15}$$

Each of these maps represent a coarsening of the partition $\{1, 1, 1\}$. A partition \mathcal{P}_1 is *coarser* than a partition \mathcal{P}_2 if \mathcal{P}_1 is made by combining elements of \mathcal{P}_2 (\mathcal{P}_2 is *finer* than \mathcal{P}_1). These maps can be used to determine all ways to rewrite a given $s(\mathcal{P})$. For a given partition \mathcal{P} , we write all maps to coarser partitions (or itself), and each step in the map is associated with a factor that either merges or clusters particles. For instance, in the map $\{1, 1, 1\} \rightarrow \{1, 2\} \rightarrow \{3\}$, the associated factors are

$$\begin{array}{rclcl}
\text{map} : & \{1, 1, 1\} & \longrightarrow & \{1, 2\} & \longrightarrow & \{3\} \\
\text{contribution} : & & -3c(2) & & -m(\{1, 2\}) &
\end{array} \tag{2.16}$$

In the first step, we cluster two single particles, and there are 3 ways to choose a pair of particles. In the second step, we merge the single particle with the clustered pair. The minus signs arise because we are using a unitarity relation, as in Eq. (2.14). Note that whenever a map involves only grouping single particles, e.g. $\{1, 1, 1\} \rightarrow \{1, 2\}$, a clustering factor is generated, and whenever it involves a group with multiple particles, e.g. $\{1, 2\} \rightarrow \{3\}$, a merging factor is generated.

We can think of a map, ϕ , as a sequence that successively coarsens the initial partition:

$$\phi(\mathcal{P}_1, \mathcal{P}_N) = \{\mathcal{P}_1, \dots, \mathcal{P}_N\}, \text{ with } \mathcal{P}_{i+1} \text{ coarser than } \mathcal{P}_i. \tag{2.17}$$

Each step along the sequence is associated with a clustering or merging factor. We define the contribution from each step in the map to not include the functions $s(\mathcal{P}_i)$ appearing in the full phase space constraints. To formalize this, we will define a few special partitions:

$$\mathcal{P}^0 : \text{the base partition } \{1, \dots, 1\}. \tag{2.18}$$

$$\mathcal{P}^w : \text{the partition that gives the outcome of the algorithm}$$

$$\mathcal{P}^f : \text{the final partition in a map } \phi.$$

In Eq. (2.8), we obtain a factor of $s(\mathcal{P}^w)$ that we want to rewrite in terms of inclusive c and m functions. The sum over all possible maps $\phi(\mathcal{P}^w, \mathcal{P}^f)$, including summing over all possible \mathcal{P}^f , gives the unitarity relation that lets us rewrite $s(\mathcal{P}^w)$. By definition, \mathcal{P}^f must be coarser than or equal to \mathcal{P}^w .

As an example, for 4 particles we can use these maps to determine the unitarity relation

for $s(\{1, 1, 1, 1\}) = s(\mathcal{P}^0)$. These maps and the associated factors are

$\phi : \mathcal{P}^w$	\longrightarrow	\mathcal{P}^f	:	contribution
\mathcal{P}^0	\longrightarrow	$\{1, 1, 1, 1\}$:	1
\mathcal{P}^0	\longrightarrow	$\{1, 1, 2\}$:	$-6 c(2)$
\mathcal{P}^0	\longrightarrow	$\{1, 3\}$:	$-4 c(3)$
$\mathcal{P}^0 \rightarrow \{1, 1, 2\}$	\rightarrow	$\{1, 3\}$:	$[-6 c(2)][-2 m(\{1, 2\})]$
\mathcal{P}^0	\longrightarrow	$\{2, 2\}$:	$-3 c(2)^2$
$\mathcal{P}^0 \rightarrow \{1, 1, 2\}$	\rightarrow	$\{2, 2\}$:	$[-6 c(2)][-c(2)]$
\mathcal{P}^0	\longrightarrow	$\{4\}$:	$-c(4)$
$\mathcal{P}^0 \rightarrow \{1, 3\}$	\rightarrow	$\{4\}$:	$[-4 c(3)][-m(\{1, 3\})]$
$\mathcal{P}^0 \rightarrow \{2, 2\}$	\rightarrow	$\{4\}$:	$[-3 c(2)^2][-m(\{2, 2\})]$
$\mathcal{P}^0 \rightarrow \{1, 1, 2\}$	\rightarrow	$\{4\}$:	$[-6 c(2)][-m(\{1, 1, 2\})]$
$\mathcal{P}^0 \rightarrow \{1, 1, 2\} \rightarrow \{1, 3\}$	\rightarrow	$\{4\}$:	$[-6 c(2)][-2 m(\{1, 2\})][-m(\{1, 3\})]$
$\mathcal{P}^0 \rightarrow \{1, 1, 2\} \rightarrow \{2, 2\}$	\rightarrow	$\{4\}$:	$[-6 c(2)][-c(2)][-m(\{2, 2\})]$

(2.19)

Summing over all maps gives the unitarity relation

$$\begin{aligned}
s(\{1, 1, 1, 1\}) &= 1 - 6 c(2) - 4 c(3) + 12 c(2) m(\{1, 2\}) - 3 c(2)^2 + 6 c(2)^2 - c(4) \\
&\quad + 4 c(3) m(\{1, 3\}) + 3 c(2)^2 m(\{2, 2\}) + 6 c(2) m(\{1, 1, 2\}) \\
&\quad - 12 c(2) m(\{1, 2\}) m(\{1, 3\}) - 6 c(2)^2 m(\{2, 2\}).
\end{aligned} \tag{2.20}$$

We give the clustering diagrams for 4 particles in Fig. 6, which can alternatively be used to derive this relation.

This map structure is very useful once we add in the measurement on the final state. There is another important identification to make in these maps. We can divide a map $\mathcal{P}^w \rightarrow \mathcal{P}^f$ into two parts:

$$\begin{array}{ccccc}
\phi : \mathcal{P}^w & \xrightarrow{\quad\quad\quad} & \mathcal{P}^c & \xrightarrow{\quad\quad\quad} & \mathcal{P}^f \\
& \text{clustering} & & \text{merging/clustering} &
\end{array} \tag{2.21}$$

as in

$$\phi(\mathcal{P}^w, \mathcal{P}^f) = \phi_w(\mathcal{P}^w, \mathcal{P}^c) \cdot \phi_c(\mathcal{P}^c, \mathcal{P}^f). \tag{2.22}$$

This will provide a key simplification in the measurement function. The map $\mathcal{P}^w \rightarrow \mathcal{P}^c$ involves only clustering functions. The map $\mathcal{P}^c \rightarrow \mathcal{P}^f$ involves merging or clustering functions, but the first step in the map must be merging so that \mathcal{P}^c is unambiguously defined. \mathcal{P}^c can be equal to \mathcal{P}^w or \mathcal{P}^f , depending on the map. To make the role of \mathcal{P}^c clear, we repeat the 4 particle maps with \mathcal{P}^c explicitly identified.

n = 4

$$\begin{aligned}
& \begin{array}{c} \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline 1 \end{array} = \begin{array}{c} \bullet \bullet \bullet \bullet \\ \hline c(4) \end{array} + \begin{array}{c} \bullet \bullet \bullet \bullet | \\ \hline 4 c(3) s(\{1, 3\}) \end{array} + \begin{array}{c} \bullet \bullet | \bullet \bullet \\ \hline 3 c(2)^2 s(\{2, 2\}) \end{array} \\
& \quad + \begin{array}{c} \bullet \bullet | \bullet | \bullet \\ \hline 6 c(2) s(\{1, 1, 2\}) \end{array} + \begin{array}{c} | \bullet | \bullet | \bullet | \bullet \\ \hline s(\{1, 1, 1, 1\}) \end{array} \\
\\
& \begin{array}{c} \bullet \bullet \bullet \\ \hline c(3) \end{array} = \begin{array}{c} \bullet \bullet \bullet | \bullet \\ \hline c(3) m(\{1, 3\}) \end{array} + \begin{array}{c} \bullet \bullet \bullet \bullet | \\ \hline c(3) s(\{1, 3\}) \end{array} \\
\\
& \begin{array}{c} \bullet \bullet \\ \hline c(2)^2 \end{array} = \begin{array}{c} \bullet \bullet | \bullet \bullet \\ \hline c(2)^2 m(\{2, 2\}) \end{array} + \begin{array}{c} \bullet \bullet | \bullet \bullet \\ \hline c(2)^2 s(\{2, 2\}) \end{array} \\
\\
& \begin{array}{c} \bullet \bullet \\ \hline c(2) \end{array} = \begin{array}{c} \bullet \bullet | \bullet \bullet | \bullet \bullet \\ \hline c(2) m(\{1, 1, 2\}) \end{array} + \begin{array}{c} \bullet \bullet | \bullet \bullet | \bullet \\ \hline 2 c(2) m(\{1, 2\}) s(\{1, 3\}) \end{array} \\
& \quad + \begin{array}{c} \bullet \bullet | \bullet | \bullet \bullet \\ \hline c(2) s(\{1, 1, 2\}) \end{array} + \begin{array}{c} \bullet \bullet | \bullet \bullet | \bullet \bullet \\ \hline c(2)^2 s(\{2, 2\}) \end{array}
\end{aligned}$$

Figure 6: Clustering diagrams for the case $n = 4$. The meaning of the diagrams are the same as in Fig. 5. These can be used to derive the unitarity relations for 4 particles, such as $s(\{1, 1, 1, 1\})$.

$$\begin{array}{ccccccc}
\phi: & \mathcal{P}^w & \longrightarrow & \mathcal{P}^c & \longrightarrow & \mathcal{P}^f & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 1, 1\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 2\} & \longrightarrow & \{1, 1, 2\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 3\} & \longrightarrow & \{1, 3\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 2\} & \longrightarrow & \{1, 3\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{2, 2\} & \longrightarrow & \{2, 2\} & \\
& \{1, 1, 1, 1\} \rightarrow \{1, 1, 2\} \rightarrow & \{2, 2\} & \longrightarrow & \{2, 2\} & (2.23) & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{4\} & \longrightarrow & \{4\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 3\} & \longrightarrow & \{4\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{2, 2\} & \longrightarrow & \{4\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 2\} & \longrightarrow & \{4\} & \\
& \{1, 1, 1, 1\} & \longrightarrow & \{1, 1, 2\} \rightarrow \{1, 3\} \rightarrow & \{4\} & & \\
& \{1, 1, 1, 1\} \rightarrow \{1, 1, 2\} \rightarrow & \{2, 2\} & \longrightarrow & \{4\} & &
\end{array}$$

Note that parts of the map where the partition does not change (such as $\{1, 1, 2\} \rightarrow \{1, 1, 2\}$) contribute nothing to the unitarity relation.

So far we have only discussed phase space constraints from clustering. We will now include the effects of the measurement, where much of this structure is employed. For reference, we give a brief dictionary of symbols defined up to this point in Table 1.

3. The Measurement Function for Clustering Algorithms

The outcome of a jet algorithm is associated with a measurement on the final state. Each partition \mathcal{P} gives a contribution $\widetilde{\mathcal{M}}(\mathcal{P})$ to the measurement function, where $\widetilde{\mathcal{M}}(\mathcal{P})$ is the Fourier transform of the measurement function. To find the total n -particle measurement function we can simply weight each term in the unitarity relation of Eq. (2.8) by the measurement factor for that partition, so that

$$\widetilde{\mathcal{M}}^{(n)} = \sum_{\mathcal{P}} N(\mathcal{P}) \left[\prod_{p \in \mathcal{P}} c(p) \right] \widetilde{\mathcal{M}}(\mathcal{P}) s(\mathcal{P}). \quad (3.1)$$

This is a more specific form of Eq. (1.7) in transform space. A similar form of this equation applies to momentum space; with the appropriate mapping $\widetilde{\mathcal{M}} \rightarrow \mathcal{M}$. The measurement function in Eq. (3.1) can be applied to any standard clustering algorithm with any observable. In order to simplify the effect of clustering on the structure of $\widetilde{\mathcal{M}}^{(n)}$ we would like to use the maps developed in Sec. 2.2 to rewrite $s(\mathcal{P})$ in terms of clustering and merging functions. In order to do this we require that the observables that we measure, $\{x\}$ in Eq. (1.3), are only a function of momenta in each subset p and do not depend on multiple subsets in \mathcal{P} . This is made explicit in Eq. (1.7), and hence in Eq. (3.1), where the phase

Symbol	Description
\mathcal{P}	Partition giving the outcome of the algorithm. See Eq. (2.4).
$c(p)$	Requires p particles to cluster. See Eq. (2.5).
$m(\{p_1, \dots, p_k\})$	Requires groups of p_1, \dots, p_k particles to merge. See Eq. (2.13).
$s(\{p_1, \dots, p_k\})$	Requires groups of p_1, \dots, p_k particles to not cluster. See Eq. (2.6).
\mathcal{P}^0	The base partition $\{1, \dots, 1\}$. See Eq. (2.18).
\mathcal{P}^w	The partition giving the outcome of the algorithm, used in maps. See Eq. (2.18).
\mathcal{P}^c	A coarser partition than \mathcal{P}^w , obtained by clustering single particles. See Eq. (2.23) for an example.
\mathcal{P}^f	The final step in a map, obtained from \mathcal{P}^c by merging or clustering. See Eq. (2.18).
$\phi_w(\mathcal{P}^w, \mathcal{P}^c)$	A map from \mathcal{P}^w to \mathcal{P}^c , used in rewriting $s(\mathcal{P}^w)$. See Eq. (2.21).
$\phi_c(\mathcal{P}^c, \mathcal{P}^f)$	A map from \mathcal{P}^c to \mathcal{P}^f , used in rewriting $s(\mathcal{P}^w)$. See Eq. (2.21).

Table 1: Dictionary of symbols defined to describe phase space constraints from clustering.

space constraints from clustering are separated from the measurement³. This means that the contribution to the measurement of a given outcome \mathcal{P} of running the jet algorithm can be written as

$$\widetilde{\mathcal{M}}(\mathcal{P}) = \prod_{p \in \mathcal{P}} \widetilde{w}(p), \quad (3.2)$$

where

$$\widetilde{w}(p) : \text{contribution of } p \text{ clustered particles to the measurement.} \quad (3.3)$$

Note that the single particle measurement function is $\widetilde{\mathcal{M}}^{(1)} = \widetilde{w}(1)$, so in the absence of clustering

$$\widetilde{\mathcal{M}}(\{1, \dots, 1\}) = \left[\widetilde{\mathcal{M}}^{(1)} \right]^n. \quad (3.4)$$

As an example of a measurement that satisfies Eq. (3.2), consider running an inclusive jet algorithm and measuring the jet mass for jets with $p_T > p_T^{\text{cut}}$. Each $p \in \mathcal{P}$ corresponds to a candidate jet produced by running the clustering algorithm, for which we measure its

³In principle, one can define observables that can only be written as a function of all the momenta in an event, and cannot be disentangled into contributions from separate regions of phase space. However, such observables tend not to arise in practical applications.

mass m_J^p (or x_J^p in Fourier space). In this case $\widetilde{\mathcal{M}}(\mathcal{P})$ becomes

$$\widetilde{\mathcal{M}}(\mathcal{P}) = \prod_{p \in \mathcal{P}} \exp \left[-i x_J^p \theta(p_T(p) > p_T^{\text{cut}}) \sum_{i,j \in p} q_i \cdot q_j \right], \quad (3.5)$$

where q_i are the momenta of particles in a given partition p . Note that for this observable, the one particle measurement function is $\widetilde{\mathcal{M}}(1) = 1$ for massless final states.

For the very general class of observables satisfying Eq. (3.2), the n -particle measurement function in Eq. (3.1) becomes

$$\widetilde{\mathcal{M}}^{(n)} = \sum_{\mathcal{P}} N(\mathcal{P}) \left[\prod_{p \in \mathcal{P}} c(p) \widetilde{w}(p) \right] s(\mathcal{P}), \quad (3.6)$$

where all partitions have n particles. We now show how the maps in Sec. 2.2 can be used to rewrite $s(\mathcal{P})$ to express the full measurement function in a form that makes the effect of clustering explicit.

We can think of the measurement as a map from the base partition $\mathcal{P}^0 = \{1, \dots, 1\}$ to the outcome of the algorithm \mathcal{P}^w :

$$\text{measurement map } \phi_m(\mathcal{P}^0, \mathcal{P}^w) : \mathcal{P}^0 \rightarrow \mathcal{P}^w. \quad (3.7)$$

Each map is associated with the clustering and measurement factor for the partition. The full phase space constraints includes a factor of $s(\mathcal{P})$, but we define the contribution from the map ϕ_m without this factor. The $s(\mathcal{P})$ factors will be rewritten using the maps ϕ_w and ϕ_c from Sec. 2.2. For 3 particles, the measurement maps and contribution to Eq. (3.6) are

$\phi_m :$	\mathcal{P}^0	\rightarrow	\mathcal{P}^w	contribution	
	$\{1, 1, 1\}$	\rightarrow	$\{1, 1, 1\}$	$:$	$\widetilde{w}(1)^3$
	$\{1, 1, 1\}$	\rightarrow	$\{3\}$	$:$	$c(3)\widetilde{w}(3)$
	$\{1, 1, 1\}$	\rightarrow	$\{1, 2\}$	$:$	$3 c(2)\widetilde{w}(2)\widetilde{w}(1)$

(3.8)

The measurement map corresponds to Eq. (3.6) with \mathcal{P} replaced by \mathcal{P}^w . The remaining work is to rewrite $s(\mathcal{P}^w)$ in terms of inclusive constraints, which we use the clustering maps for.

Putting the measurement map together with the clustering maps defined in Sec. 2.2, we have a complete map from the base partition to the final partition:

$$\phi_{\text{tot}}(\mathcal{P}^0, \mathcal{P}^f) = \phi_m(\mathcal{P}^0, \mathcal{P}^w) \cdot \phi_w(\mathcal{P}^w, \mathcal{P}^c) \cdot \phi_c(\mathcal{P}^c, \mathcal{P}^f). \quad (3.9)$$

If we write all such maps, then we can determine the outcome of the algorithm.

As an example, we give the measurement functions for the 2 and 3 particle cases. In Fig. 7 we use the diagrammatic notation of Sec. 2 to show the measurement maps $\mathcal{P}^0 \rightarrow \mathcal{P}^w$ and clustering maps $\mathcal{P}^w \rightarrow \mathcal{P}^c \rightarrow \mathcal{P}^f$ for the 3 particle case. For $n = 2$, there are three

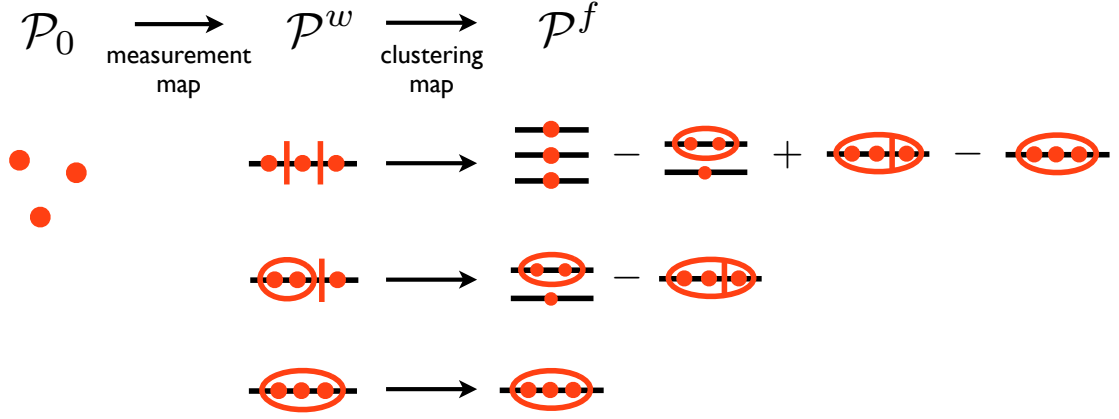


Figure 7: The maps to determine the measurement function for 3 particles are shown diagrammatically. The measurement map $\mathcal{P}^0 \rightarrow \mathcal{P}^w$ corresponds to Eq. (3.8), while the clustering maps $\mathcal{P}^w \rightarrow \mathcal{P}^f$ rewrite $s(\mathcal{P}^w)$ using the unitarity relations displayed in Fig. 5. The contribution from each map is given in Eq. (3.12).

maps.

$\phi_{\text{tot}} :$	\mathcal{P}^0	\rightarrow	\mathcal{P}^w	\rightarrow	\mathcal{P}^c	\rightarrow	\mathcal{P}^f	$:$	contribution	
	$\{1, 1\}$	\rightarrow	$\{1, 1\}$	\rightarrow	$\{1, 1\}$	\rightarrow	$\{1, 1\}$	$:$	$\tilde{w}(1)^2$	(3.10)
	$\{1, 1\}$	\rightarrow	$\{1, 1\}$	\rightarrow	$\{2\}$	\rightarrow	$\{2\}$	$:$	$-c(2)\tilde{w}(1)^2$	
	$\{1, 1\}$	\rightarrow	$\{2\}$	\rightarrow	$\{2\}$	\rightarrow	$\{2\}$	$:$	$c(2)\tilde{w}(2)$	

Therefore the total measurement function is

$$\begin{aligned}
 \widetilde{\mathcal{M}}^{(2)} &= \tilde{w}(1)^2 + c(2)[\tilde{w}(2) - \tilde{w}(1)^2] \\
 &\equiv [\widetilde{\mathcal{M}}^{(1)}]^2 + \Delta\widetilde{\mathcal{M}}^{(2)},
 \end{aligned} \tag{3.11}$$

which defines $\Delta\widetilde{\mathcal{M}}^{(2)}$. For $n = 3$, when the measurement maps from Eq. (3.8) are combined with clustering maps, which rewrite $s(\mathcal{P}^w)$, there are seven maps.

$\phi_{\text{tot}} :$	\mathcal{P}^0	\rightarrow	\mathcal{P}^w	\rightarrow	\mathcal{P}^c	\rightarrow	\mathcal{P}^f	$:$	contribution
	\mathcal{P}^0	\rightarrow	$\{1, 1, 1\}$	\rightarrow	$\{1, 1, 1\}$	\rightarrow	$\{1, 1, 1\}$	$:$	$\tilde{w}(1)^3$
	\mathcal{P}^0	\rightarrow	$\{1, 1, 1\}$	\rightarrow	$\{1, 2\}$	\rightarrow	$\{1, 2\}$	$:$	$[\tilde{w}(1)^3][-3c(2)]$
	\mathcal{P}^0	\rightarrow	$\{1, 2\}$	\rightarrow	$\{1, 2\}$	\rightarrow	$\{1, 2\}$	$:$	$3c(2)\tilde{w}(2)\tilde{w}(1)$
	\mathcal{P}^0	\rightarrow	$\{1, 1, 1\}$	\rightarrow	$\{1, 2\}$	\rightarrow	$\{3\}$	$:$	$[\tilde{w}(1)^3][-3c(2)][-m(\{1, 2\})]$
	\mathcal{P}^0	\rightarrow	$\{1, 2\}$	\rightarrow	$\{1, 2\}$	\rightarrow	$\{3\}$	$:$	$[3c(2)\tilde{w}(2)\tilde{w}(1)][-m(\{1, 2\})]$
	\mathcal{P}^0	\rightarrow	$\{1, 1, 1\}$	\rightarrow	$\{3\}$	\rightarrow	$\{3\}$	$:$	$[\tilde{w}(1)^3][-c(3)]$
	\mathcal{P}^0	\rightarrow	$\{3\}$	\rightarrow	$\{3\}$	\rightarrow	$\{3\}$	$:$	$c(3)\tilde{w}(3)$

(3.12)

These combined to give the measurement function

$$\begin{aligned}\widetilde{\mathcal{M}}^{(3)} &= \widetilde{w}(1)^3 + 3c(2)[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1) \\ &\quad + \left(c(3)[\widetilde{w}(3) - \widetilde{w}(1)^3] - 3c(2)m(\{1, 2\})[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1)\right) \\ &\equiv [\widetilde{\mathcal{M}}^{(1)}]^3 + 3\widetilde{\mathcal{M}}^{(1)}\Delta\widetilde{\mathcal{M}}^{(2)} + \Delta\widetilde{\mathcal{M}}^{(3)}.\end{aligned}\tag{3.13}$$

This equation defines $\Delta\widetilde{\mathcal{M}}^{(3)}$. Note that for the 2 and 3 particle cases, a factor of $\widetilde{w}(k)$ for $k > 1$ comes with an accompanying $-\widetilde{w}(1)^k$. This is true in general, which we can show by making use of the separation between the maps ϕ_w and ϕ_c . This will lead to a simplification of the structure of the measurement function.

Consider the set of maps $\phi_m(\mathcal{P}^0, \mathcal{P}^w) \cdot \phi_w(\mathcal{P}^w, \mathcal{P}^c)$ from \mathcal{P}^0 to \mathcal{P}^c with a fixed \mathcal{P}^c . These maps are generated by taking all possible \mathcal{P}^w . For example, for $n = 5$ and $\mathcal{P}^c = \{1, 2, 2\}$, these maps are (with \mathcal{P}^w in bold for each map):

$\phi_m \cdot \phi_w :$	\mathcal{P}^0	\longrightarrow	\mathcal{P}^c	$:$	contribution
	$\{\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}\}$	\rightarrow	$\{1, 1, 1, 2\}$	\rightarrow	$\{1, 2, 2\} : 15 c(2)^2 \widetilde{w}(1)^5$
	$\{1, 1, 1, 1, 1\}$	\rightarrow	$\{\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{2}\}$	\rightarrow	$\{1, 2, 2\} : -30 c(2)^2 \widetilde{w}(2) \widetilde{w}(1)^3$
	$\{1, 1, 1, 1, 1\}$	\rightarrow	$\{1, 1, 1, 2\}$	\rightarrow	$\{\mathbf{1}, \mathbf{2}, \mathbf{2}\} : 15 c(2)^2 \widetilde{w}(2)^2 \widetilde{w}(1)$

(3.14)

The sum of these terms is

$$15 c(2)^2 [\widetilde{w}(2) - \widetilde{w}(1)^2]^2 \widetilde{w}(1).\tag{3.15}$$

This binomial structure persists. All terms will share the same clustering factors, differing only in the combinatoric and measurement factors. If we sum over all \mathcal{P}^w for a given \mathcal{P}^c ($\mathcal{P}^w | \mathcal{P}^c$), then the clustering and measurement factors are

$$N(\mathcal{P}^c) \left[\prod_{p \in \mathcal{P}^c} c(p) \right] \sum_{\mathcal{P}^w | \mathcal{P}^c} N_w(\mathcal{P}^w, \mathcal{P}^c) \left[\prod_{p \in \mathcal{P}^w} \widetilde{w}(p) \right].\tag{3.16}$$

The combinatoric factor is defined relative to $N(\mathcal{P}^c)$, and so it is the number of ways to distinguish the map $\mathcal{P}^0 \rightarrow \mathcal{P}^w \rightarrow \mathcal{P}^c$ from $\mathcal{P}^0 \rightarrow \mathcal{P}^c$. This is simply

$$N_w(\mathcal{P}^w, \mathcal{P}^c) = \prod_{k=2}^n \binom{d_k^c}{d_k^w},\tag{3.17}$$

where d_k^i is the number of elements of \mathcal{P}^i equal to k (recall Eq. (2.3)). Therefore, we can sum over all \mathcal{P}^w and Eq. (3.16) becomes

$$N(\mathcal{P}^c) \left[\prod_{p \in \mathcal{P}^c} c(p) \Delta \widetilde{w}(p) \right],\tag{3.18}$$

where

$$\Delta \widetilde{w}(p) = \begin{cases} \widetilde{w}(p) - \widetilde{w}(1)^p & \text{if } p > 1, \\ \widetilde{w}(1) & \text{if } p = 1. \end{cases}\tag{3.19}$$

Finally, we must deal with the map $\phi_c(\mathcal{P}^c, \mathcal{P}^f)$. Given a partition \mathcal{P}^c , we will define the set of all maps $\phi_c(\mathcal{P}^c, \mathcal{P}^f)$ (all such maps for all allowed \mathcal{P}^f) as $\mathcal{G}_c(\mathcal{P}^c)$. For example,

$$\mathcal{G}_c(\{1, 1, 2\}) = \{\phi_c(\{1, 1, 2\}, \{1, 1, 2\}), \phi_c(\{1, 1, 2\}, \{1, 3\}), \phi_c(\{1, 1, 2\}, \{4\})\}. \quad (3.20)$$

These maps from \mathcal{P}^c to \mathcal{P}^f are the second, third, and eighth maps in Eq. (2.23). For each map ϕ_c , we define

$$\mathcal{F}(\phi_c) : \text{the product of inclusive constraints for the map } \phi_c. \quad (3.21)$$

Using the maps in $\mathcal{G}_c(\{1, 1, 2\})$ as an example,

$$\begin{aligned} \mathcal{F}(\phi_c(\{1, 1, 2\}, \{1, 1, 2\})) &= 1 \\ \mathcal{F}(\phi_c(\{1, 1, 2\}, \{1, 3\})) &= -2 m(\{1, 2\}) \\ \mathcal{F}(\phi_c(\{1, 1, 2\}, \{4\})) &= 2 m(\{1, 2\}) m(\{1, 3\}). \end{aligned} \quad (3.22)$$

3.1 Main Results

Putting the pieces together, the measurement function is

$$\boxed{\mathcal{M}^{(n)} = \sum_{\mathcal{P}^c} N(\mathcal{P}^c) \left[\prod_{p \in \mathcal{P}^c} c(p) \Delta \tilde{w}(p) \right] \sum_{\phi \in \mathcal{G}_c(\mathcal{P}^c)} \mathcal{F}(\phi).} \quad (3.23)$$

This is one of our main results, and can be used to arrange the measurement function into a form suitable to study the logarithmic contributions to the cross section.

Each map $\phi : \mathcal{P}^0 \rightarrow \mathcal{P}^f$ will contain some number of single particles not participating in any clustering or merging. These particles will only be associated with factors of $\tilde{w}(1) = \tilde{\mathcal{M}}^{(1)}$. Since $\tilde{\mathcal{M}}^{(1)}$ only depends on the phase space constraints from a single particle, it is useful to distinguish these contributions from multi-particle phase space constraints⁴.

Suppose we consider maps with exactly k factors of $\tilde{\mathcal{M}}^{(1)}$. These single particles do not participate in the maps at all, and will remain in the final partition \mathcal{P}^f for each map. Therefore we can factor out these particles from the map, and they will become maps on $n - k$ particles:

$$\phi_n \rightarrow \phi_{n-k} \quad (3.24)$$

If the map ϕ_n leads to a contribution $\tilde{\mathcal{M}}(\phi_n)$ to the measurement function, then we can write this as

$$\tilde{\mathcal{M}}(\phi_n) = \frac{n!}{(n-k)! k!} \tilde{\mathcal{M}}(\phi_{n-k}) [\tilde{\mathcal{M}}^{(1)}]^k. \quad (3.25)$$

The combinatoric factor changes the $1/n!$ from the n -particle matrix element into separate factors for $\tilde{\mathcal{M}}(\phi_{n-k})$ and $[\tilde{\mathcal{M}}^{(1)}]^k$. This gives another of our main results,

$$\boxed{\tilde{\mathcal{M}}^{(n)} = [\tilde{\mathcal{M}}^{(1)}]^n + \sum_{k=0}^{n-2} \frac{n!}{(n-k)! k!} \Delta \tilde{\mathcal{M}}^{(n-k)} [\tilde{\mathcal{M}}^{(1)}]^k,} \quad (3.26)$$

⁴When we consider boundary clustering, in many cases $\tilde{\mathcal{M}}^{(1)}$ is the only factor that contributes at leading log to the cross section. For an example, see [27], where the contribution $\Delta \tilde{\mathcal{M}}^{(2)}$ was calculated for the C/A and k_T algorithms for a jet mass observable in dijet events. This will be discussed further in Sec. 5.

where $\Delta\widetilde{\mathcal{M}}^{(n)}$ are the n -particle measurement function terms where all particles cluster or merge with at least one other particle, and we can use this equation to define $\Delta\widetilde{\mathcal{M}}^{(n)}$. For 2 particles,

$$\begin{aligned}\Delta\widetilde{\mathcal{M}}^{(2)} &= c(2)[\widetilde{w}(2) - \widetilde{w}(1)^2] \\ &= c(2)\Delta\widetilde{w}(2),\end{aligned}\tag{3.27}$$

and for 3 particles,

$$\begin{aligned}\Delta\widetilde{\mathcal{M}}^{(3)} &= c(3)[\widetilde{w}(3) - \widetilde{w}(1)^3] - 3c(2)m(\{1, 2\})[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1) \\ &= c(3)\Delta\widetilde{w}(3) - 3c(2)m(\{1, 2\})\Delta\widetilde{w}(2)\widetilde{w}(1).\end{aligned}\tag{3.28}$$

We give the measurement functions for $n = 4$ and $n = 5$ in Appendix A as additional examples.

4. Cancellation of Infrared Divergences from Clustering

The formalism for the measurement function lets us study the IR structure of the various clustering contributions for the Abelian terms. As an example, we consider the 3 particle terms in the measurement function in the limit that one particle is very soft. This argument can be generalized to higher order terms, or adapted to the collinear limit. For 3 particles,

$$\begin{aligned}\widetilde{\mathcal{M}}^{(3)} &= [\widetilde{\mathcal{M}}^{(1)}]^3 + 3\Delta\widetilde{\mathcal{M}}^{(2)}\widetilde{\mathcal{M}}^{(1)} + \Delta\widetilde{\mathcal{M}}^{(3)} \\ &= \widetilde{w}(1)^3 + 3c(2)[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1) \\ &\quad + c(3)[\widetilde{w}(3) - \widetilde{w}(1)^3] - 3c(2)m(\{1, 2\})[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1).\end{aligned}\tag{4.1}$$

The single particle measurement function $\widetilde{\mathcal{M}}^{(1)}$ leads to an IR safe contribution to the cross section. We will consider $\Delta\widetilde{\mathcal{M}}^{(3)}$, which is the last line of Eq. (4.1); in showing these terms are IR safe, we will show that terms involving $\Delta\widetilde{\mathcal{M}}^{(2)}$ are also free of IR divergences.

Let us label the soft momenta k_1 , k_2 , and k_3 ; we will take the limit $k_3 \rightarrow 0$. The term involving the constraints $c(2)m(\{1, 2\})$ has been symmetrized over soft momenta, meaning if we want to take k_3 soft then we must split this up into 3 separate terms. Labeling the momenta, this term becomes

$$\begin{aligned}&3c(2)m(\{1, 2\})[\widetilde{w}(2) - \widetilde{w}(1)^2]\widetilde{w}(1) \\ &\rightarrow c(\{k_1, k_2\})m(\{k_3, \{k_1, k_2\}\})[\widetilde{w}(k_1, k_2) - \widetilde{w}(k_1)\widetilde{w}(k_2)]\widetilde{w}(k_3) \\ &\quad + c(\{k_1, k_3\})m(\{k_2, \{k_1, k_3\}\})[\widetilde{w}(k_1, k_3) - \widetilde{w}(k_1)\widetilde{w}(k_3)]\widetilde{w}(k_2) \\ &\quad + c(\{k_2, k_3\})m(\{k_1, \{k_2, k_3\}\})[\widetilde{w}(k_2, k_3) - \widetilde{w}(k_2)\widetilde{w}(k_3)]\widetilde{w}(k_1).\end{aligned}\tag{4.2}$$

In this language, the IR safety of $\Delta\widetilde{\mathcal{M}}^{(2)}$ is due to the relations

$$\lim_{k_s \rightarrow 0} \widetilde{w}(k, k_s) = \widetilde{w}(k), \quad \lim_{k_s \rightarrow 0} \widetilde{w}(k_s) = 1.\tag{4.3}$$

This implies that 2 particle clustering effects are IR safe:

$$\lim_{k_s \rightarrow 0} \widetilde{w}(k, k_s) - \widetilde{w}(k)\widetilde{w}(k_s) = 0 \quad \Rightarrow \quad \lim_{k_s \rightarrow 0} \Delta\widetilde{\mathcal{M}}^{(2)} = 0.\tag{4.4}$$

This relation simplifies the terms in Eq. (4.2) to

$$c(\{k_1, k_2\})m(\{k_3, \{k_1, k_2\}\})[\tilde{w}(k_1, k_2) - \tilde{w}(k_1)\tilde{w}(k_2)] . \quad (4.5)$$

The relation in Eq. (4.3) can be generalized to

$$\lim_{k_s \rightarrow 0} \tilde{w}(n) \rightarrow \tilde{w}(n-1) , \quad (4.6)$$

which means that the $c(3)$ term in $\Delta\widetilde{\mathcal{M}}^{(3)}$ becomes

$$c(\{k_1, k_2, k_3\})[\tilde{w}(\{k_1, k_2\}) - \tilde{w}(k_1)\tilde{w}(k_2)] . \quad (4.7)$$

Therefore both groups of terms are IR divergent in the soft limit. However, by considering the merging constraints one can show for an IR safe jet algorithm that

$$\lim_{k_3 \rightarrow 0} c(\{k_1, k_2\})m(\{k_3, \{k_1, k_2\}\}) = c(\{k_1, k_2, k_3\}) . \quad (4.8)$$

The reason is that when k_3 is soft, clustering with any other particle k_i would leave the momentum of k_i unchanged. In this limit the constraints requiring the soft particle to merge with a clustered pair is the same as all three particles clustering. Therefore, both terms in $\Delta\widetilde{\mathcal{M}}^{(3)}$ reduce to Eq. (4.7), and they cancel ($\Delta\widetilde{\mathcal{M}}^{(3)} \rightarrow 0$). This is an explicit demonstration that IR divergences are not introduced by clustering effects for 2 or 3 particles.

5. An Application: All-Orders Abelian Structure of Soft Clustering for Jet Shapes

The all-orders description of the measurement function developed in the previous sections is useful because it enables us to understand the effects of clustering on the higher order terms in the cross section for jet observables. As an example we consider $e^+e^- \rightarrow N$ jets defined by an inclusive jet algorithm (such as C/A) and define two observables, ρ and Λ , where $\rho = (m_1^2 + \dots + m_N^2)/Q^2$ is the sum of jet masses $\{m_i\}$ scaled by the center of mass energy Q and Λ is the total energy of radiation outside of all jets. When $\{Q\rho, \Lambda\} \ll Q\sqrt{\rho} \ll Q$, each jet mass m_i is forced to be small and the event is described by collinear radiation in the jets and soft radiation in and between the jets. This observable was considered for dijet events in [30, 26, 27, 31, 32, 33]. The all-orders structure of the Abelian terms in this section apply to a wide range of jet shapes in the soft/collinear regime, such as angularities.

Perturbative calculations generate logarithms of the form $\ln \rho$ and $\ln \Lambda/Q$, and a reliable prediction in this regime requires resummation. This is achieved through a factorization theorem, for which for the double cumulant takes the form

$$\Sigma(\rho_c, \Lambda_c) = \int_0^{\rho_c} d\rho \int_0^{\Lambda_c} d\Lambda \frac{d^2\sigma}{d\rho d\Lambda} , \quad (5.1)$$

where

$$\frac{d^2\sigma}{d\rho d\Lambda} = \sigma_0 H_N(Q) \left[J_1(\rho) \otimes \dots \otimes J_N(\rho) \right] \otimes S_N(\rho, \Lambda) . \quad (5.2)$$

The hard function H is determined by the short distance interaction that produces the high energy jets, and depends on the center of mass energy Q . The jet functions parameterize collinear radiation for each jet at the scale $Q\sqrt{\rho}$ and the soft function parameterizes soft radiation in and between jets at the scale $Q\rho$. The radiation collinear to each jet direction clusters only within the jet when $\sqrt{\rho} \ll R$, and in this case the effect of clustering is described entirely by the *boundary clustering* of soft particles in the event, as discussed in Sec. 1. It was shown in [26, 27] that boundary clustering gives rise to clustering NGLs in the soft function of the form $\ln(Q\rho/\Lambda)$ starting at $\mathcal{O}(\alpha_s^2 \ln^2)$.

$$S_N(\rho, \Lambda) = \frac{1}{c_N} \langle 0 | \widehat{\mathcal{W}}_{\{n\}}^\dagger \widehat{\mathcal{M}}(\{\rho, \Lambda\}) \widehat{\mathcal{W}}_{\{n\}} | 0 \rangle \quad (5.3)$$

where $\widehat{\mathcal{W}}_{\{n\}}$ represents a product of soft Wilson lines, Y_{n_i} , where $n_i^\mu = (1, \hat{n}_i)$ is a light like vector in the i -th jet direction, with the appropriate color representation and path ordering of the final state hard partons⁵. The soft function is a matrix in color space and c_N is the color normalization factor such that $S_N(\rho, \Lambda) = \delta(\rho)\delta(\Lambda)$ at tree-level. The measurement operator $\widehat{\mathcal{M}}(\rho, \Lambda)$ implements the measurement of ρ and Λ for a given jet algorithm, and, for final states $|X_s\rangle$ with soft momenta $\{k\}$, the measurement function acts as

$$\widehat{\mathcal{M}}(\rho, \Lambda) |X_s\rangle = \mathcal{M}(\rho, \Lambda, \{k\}) |X_s\rangle. \quad (5.4)$$

For convenience we work with the Fourier space soft function,

$$\widetilde{S}_N(x, y) = \int_{-\infty}^{\infty} d\rho d\Lambda \exp(-ix\rho) \exp(-iy\Lambda/Q) S_N(\rho, \Lambda), \quad (5.5)$$

for which the measurement function is typically expressed in the form

$$\widetilde{\mathcal{M}}(x, y, \{k\}) = \prod_{i=1}^N \exp\left(-ix \sum_j \frac{n_i \cdot k_j}{Q} \mathcal{R}_{J_i}(k_j; \{k\})\right) \exp\left(-iy \sum_j \frac{k_j^0}{Q} \mathcal{R}_{\text{out}}(k_j; \{k\})\right), \quad (5.6)$$

where the functions $\mathcal{R}_{J_i}(k_j; \{k\})$ and $\mathcal{R}_{\text{out}}(k_j; \{k\})$ implement the phase space constraints of the jet algorithm, requiring particle k_j to be in the i -th jet and out-of-jet regions respectively. Each region, \mathcal{R}_i , depends on all the soft momenta in the final state $\{k\}$ and therefore implicitly on all of the other regions. The formalism developed in this paper helps disentangle this dependence. For n final state particles it is useful to organize the effects of clustering as in Eq. (1.13). When the measurement function is integrated against the squared matrix element from the Wilson lines, we get the $\mathcal{O}(\alpha_s^n)$ contribution to the soft function,

$$\widetilde{S}^{(n)}(\{x\}) = \int \left(\prod_{i=1}^n \frac{d^4 k_i}{(2\pi)^4} \right) \mathcal{A}^{(n)}(\{k\}) \widetilde{\mathcal{M}}^{(n)}(\{x\}, \{k\}). \quad (5.7)$$

⁵Recall a soft Wilson line directed along n^μ , in representation \mathbf{T}^a , contains soft gauge fields (A_s^μ) and is defined by a path ordered exponential:

$$Y_n^\dagger(x) = \text{P exp} \left(ig \int_0^\infty ds \, n \cdot A_s^a(x + sn) \, \mathbf{T}^a \right).$$

The first term in Eq. (1.13) is the measurement function for an algorithm for which the measurement function factorizes into single particle measurements in the soft sector, such as for the anti- k_T algorithm. The second term accounts for correlations in the measurement function. This leads us to write

$$\tilde{S}^{(n)}(\{x\}) = \tilde{S}_{\text{akT}}^{(n)}(\{x\}) + \tilde{S}_{\text{corr.}}^{(n)}(\{x\}). \quad (5.8)$$

The algorithm-dependent effects of soft clustering are entirely contained in the last term.

When soft gluons boundary cluster they can not be collinear to the jet direction, as depicted in Fig. 3. In general a final state gluon can contribute a double log to the cumulant cross section, $\alpha_s \ln^2(Q\rho_c/\Lambda_c)$, associated with soft and collinear divergences. For the Abelian terms in the soft matrix element, collinear divergences arise only for soft gluons collinear to the jet direction. Therefore, as shown in [27], those gluons involved in boundary clustering can each contribute at most a single log, $\alpha_s \ln(Q\rho_c/\Lambda_c)$, from a soft divergence to the Abelian terms in $\tilde{S}_{\text{corr.}}^{(n)}$. For example, at $\mathcal{O}(\alpha_s^3)$, $\tilde{\mathcal{M}}_{\text{corr.}}^{(3)}$ contains a contribution where all three particles boundary cluster together, which contribute at NLL to $\tilde{S}_{\text{corr.}}^{(3)}$. However the measurement function also contains the contribution

$$\begin{aligned} \tilde{\mathcal{M}}^{(3)} \supset & \theta(\text{gluon 1 and 2 boundary cluster}) \\ & \times \theta(\text{gluon 3 is in the jet and does not boundary cluster}) , \end{aligned} \quad (5.9)$$

which a priori could contribute at LL in the exponent, $\alpha_s^n \ln^m$ for $m > n$, since gluon 3 is sensitive to both soft and collinear divergences. In order to show that clustering NGLs do in fact occur at NLL, we must show that constraints of the form Eq. (5.9) can be expressed in terms of lower order phase space constraints. Rewriting the measurement function in terms of inclusive constraints allows us to do precisely this. For the term in Eq. (5.9) this is illustrated graphically in Fig. 2. Next we will use our main result, Eq. (3.26), to show the all-orders form of the Abelian soft function. We will see that clustering NGLs give new contributions at NLL at all orders, spoiling the exponentiation of the Abelian terms.

5.1 All-Orders Abelian Structure

A considerable simplification happens in the Abelian sector. It is well known that the $\mathcal{O}(\alpha_s^n)$ Abelian matrix element factorizes as

$$\mathcal{A}^{(n)}(\{k_s\}) = \frac{1}{n!} \prod_{i=1}^n \mathcal{A}^{(1)}(k_i), \quad (5.10)$$

where $\mathcal{A}^{(1)}(k)$ is the squared matrix element at $\mathcal{O}(\alpha_s)$. When clustering effects are absent or power suppressed, the measurement function factorizes into the product of single gluon constraints, as for the anti- k_T algorithm in Eq. (1.11). In this case it is easy to see that

Abelian soft function exponentiates. The $\mathcal{O}(\alpha_s^n C_F^n)$ contribution to the soft function is⁶

$$\begin{aligned}\widetilde{S}_{\text{Abel.}}^{(n)}(\{x\}) &= \frac{1}{n!} \left(\int \frac{d^4 k}{(2\pi)^4} \mathcal{A}^{(1)}(k) \widetilde{\mathcal{M}}^{(1)}(\{x\}) \right)^n \\ &= \frac{1}{n!} \left[\widetilde{S}^{(1)}(\{x\}) \right]^n,\end{aligned}\tag{5.11}$$

which is summed to produces the all orders Abelian soft function,

$$\widetilde{S}_{\text{Abel.}}(\{x\}) = \exp \left[\widetilde{S}^{(1)}(\{x\}) \right].\tag{5.12}$$

This is a manifestation of Abelian exponentiation and we see that factorization of the measurement function, as in Eq. (1.11), is a necessary condition for it to hold.

For algorithms for which the measurement function contains corrections from clustering, we can use our main result, Eq. (3.26), to express these corrections as

$$\widetilde{\mathcal{M}}_{\text{corr.}}^{(n)}(\{x\}, \Phi) = \sum_{k=0}^{n-2} \frac{n!}{(n-k)! k!} \Delta \widetilde{\mathcal{M}}^{(n-k)} [\widetilde{\mathcal{M}}^{(1)}]^k,\tag{5.13}$$

where $\Delta \widetilde{\mathcal{M}}^{(n)}$ is the part of the n -particle measurement function in which all n particles cluster with at least one other particle. These corrections modify Eq. (5.11) to give

$$\widetilde{S}_{\text{Abel.}}^{(n)} = \frac{1}{n!} (\widetilde{S}^{(1)})^n + \sum_{k=0}^{n-2} \Delta \widetilde{S}_{\text{alg}}^{(n-k)} \frac{1}{k!} (\widetilde{S}^{(1)})^k,\tag{5.14}$$

where we have defined

$$\Delta \widetilde{S}_{\text{alg}}^{(n)}(\{x\}) = \frac{1}{n!} \prod_{i=1}^n \int \frac{d^4 k_i}{(2\pi)^4} \mathcal{A}^{(1)}(k_i) \Delta \widetilde{\mathcal{M}}^{(n)}(\{x\}, \{k\}).\tag{5.15}$$

Finally, although the Abelian soft function no longer exponentiates, the form of Eq. (5.14) allows us to write the all orders soft function as

$$\widetilde{S}_{\text{Abel.}}(\{x\}) = \exp \left[\widetilde{S}^{(1)}(\{x\}) \right] \left(1 + \sum_{k=2}^{\infty} \Delta \widetilde{S}_{\text{alg}}^{(k)}(\{x\}) \right).\tag{5.16}$$

Since $\Delta \widetilde{S}_{\text{alg}}^{(k)}$ arises from phase space constraints where each of the k gluons boundary clusters with *at least* one other gluon, it contributes at most at NLL; that is $\alpha_s^k \ln^k(Q\rho_c/\Lambda)$ in the cumulant cross section. Resummation of clustering NGLs would require a relation between $\Delta \widetilde{S}_{\text{alg}}^{(n)}$ for all n . However, as we have shown from the general form the measurement function in Eq. (3.23) and Eq. (3.26), at each order in α_s^n the clustering terms in the soft function receive a contribution from

$$\Delta \widetilde{\mathcal{M}}^{(n)} \supset c(n) [\widetilde{w}(n) - \widetilde{w}(1)^n],\tag{5.17}$$

where all n particles cluster with each other. This contribution is new and unique at this order in α_s . This makes resummation of clustering NGLs very unlikely. In the next subsection, we will see that the all-orders form of the Abelian soft function in Eq. (5.16) is required in order to have a renormalizable soft function.

⁶In pure dimensional regularization, Abelian matrix elements with a virtual gluon are scaleless, hence we only consider the matrix element where every particle is in the final state.

5.2 Renormalization of the Abelian Soft Function

It is well known that the renormalization of the product of Wilson lines, such as those that appear in the soft function, obey Abelian exponentiation - that is, the Abelian contribution to the all-orders anomalous dimension depends only on the one-loop result. However, the presence of the measurement operator in the soft function makes its structure and renormalization more complex. We now consider the effects that clustering has on the all orders UV structure of the Abelian soft function.

We have seen in Eq. (5.12) that since the anti- k_T measurement function can be expressed as the product of single gluon constraints, the soft function in this case obeys Abelian exponentiation. This means that the bare anti- k_T soft function is related to the renormalized one to all orders as follows:

$$\tilde{S}_{\text{ak}_T}^b = \tilde{Z}_{\text{ak}_T} \tilde{S}_{\text{ak}_T}^r = \exp \left[\tilde{Z}_{\text{ak}_T}^{(1)} + \tilde{S}^{r(1)} \right], \quad (5.18)$$

where b stands for bare, r stands for renormalized, and $\tilde{Z}_{\text{ak}_T}^{(1)}$ is the Abelian position space one-loop counter-term. We have suppressed the subscript ‘Abel.’ since all of our statements in this section only refer to Abelian terms. Consistency of the factorization in Eq. (5.2) and the requirement that the cross section must be renormalization scale independent imply that the anomalous dimensions must satisfy

$$\gamma_H + \sum_{i=1}^N \gamma_{J_i} + \gamma_S = 0. \quad (5.19)$$

Since the hard and jet function anomalous dimensions do not depend on the algorithm, the soft function anomalous dimension must also be independent of it. This implies that $\tilde{Z}_{\text{alg}}^{(1)} = \tilde{Z}_{\text{ak}_T}^{(1)}$ must hold for all algorithms. Any contribution to the soft function $\Delta \tilde{S}_{\text{alg}}$ must be UV finite, since no new divergences are introduced by these algorithms when compared to anti- k_T .

Since algorithms in the k_T class share the same \tilde{Z} factor, the bare soft function is related to the renormalized one as follows:

$$\tilde{S}_{\text{alg}}^b = \tilde{Z}_{\text{ak}_T} \tilde{S}_{\text{alg}}^r = \exp \left[\tilde{Z}_{\text{ak}_T}^{(1)} \right] \tilde{S}_{\text{alg}}^r, \quad (5.20)$$

using Eq. (5.18). When $\tilde{S}^{(1)}$ is replaced with its bare version in Eq. (5.14), it takes the form

$$\tilde{S}_{\text{alg}}^{b(n)} = \frac{1}{n!} (\tilde{S}^{b(1)})^n + \sum_{k=0}^{n-2} \Delta \tilde{S}^{(n-k)} \frac{(\tilde{S}^{b(1)})^k}{k!}. \quad (5.21)$$

We can rewrite this to all orders as

$$\tilde{S}_{\text{alg}}^b = \exp \left[\tilde{S}^{b(1)} \right] \left(1 + \sum_{k=2}^{\infty} \Delta \tilde{S}_{\text{alg}}^{(k)} \right). \quad (5.22)$$

After inserting $\tilde{S}^{b(1)} = \tilde{Z}_{\text{ak}_T}^{(1)} + \tilde{S}^{r(1)}$ into the previous equation, it follows from Eq. (5.20), that the renormalized soft function is

$$\tilde{S}_{\text{alg}}^r = \exp \left[\tilde{S}^{r(1)} \right] \left(1 + \sum_{k=2}^{\infty} \Delta \tilde{S}_{\text{alg}}^{(k)} \right). \quad (5.23)$$

Lastly, we also know from Sec. 4 that $\Delta\tilde{S}_{\text{alg}}^{(n)}$ is IR finite in addition to UV finite. As such, these terms can be calculated numerically without the need for difficult subtraction schemes. This is a major advantage, since the phase space constraints imposed by $\Delta\tilde{\mathcal{M}}^{(n)}$ are typically quite complex, making the effects difficult to calculate analytically. The $\mathcal{O}(\alpha_s^2)$ corrections for the C/A and k_T algorithms were calculated in [27] for the dijet case ($N = 2$).

5.3 Failure of Exponentiation

We might hope that the corrections due to clustering in the measurement function could be redefined to reveal further structure in the measurement function beyond that in Eq. (3.26). In particular, one might expect that the corrections to the measurement function could be redefined in such a way that the Abelian corrections exponentiate:

$$\tilde{S}_{\text{alg}}^r = \exp \left[\tilde{S}^{r(1)} + \sum_{n=2}^{\infty} \Delta\tilde{S}'^{(n)} \right]. \quad (5.24)$$

This is equivalent to reorganizing the measurement function to define a new correction term, $\Delta\tilde{\mathcal{M}}'^{(n)}$ that is integrated against the Abelian matrix element to get $\Delta\tilde{S}'^{(n)}$, as in Eq. (5.15). The form in Eq. (5.24) follows from the properties of clustering in the measurement function if we can define $\Delta\tilde{S}'^{(n)}$ at each order to involve only the clustering/merging of all n particles with each other. This is contrast with the form in Eq. (5.23), where $\Delta\tilde{S}^{(n)}$ involves clustering effects of n or fewer particles. Of course, we could simply define $\Delta\tilde{S}'$ through Eq. (5.24) at each order, since the correction terms $\Delta\tilde{S}_{\text{alg}}^{(k)}$ in Eq. (5.23) are finite. However, unless we show Eq. (5.24) from the properties of clustering in the measurement function, it has no more content than Eq. (5.23).

Terms grouping all n particles together arise, in the language of Sec. 3, from mappings that have $\mathcal{P}^f = \{n\}$. To illustrate the difference with an explicit example, consider $n = 4$, which is worked out in full detail in Appendix A. $\Delta\tilde{S}^{(4)}$ contains clustering effects of 4 particles as well as clustering effects from 2 particles in two separate groups:

$$\Delta\tilde{\mathcal{M}}^{(4)} = 3[\Delta\tilde{\mathcal{M}}^{(2)}]^2 + \Delta\tilde{\mathcal{M}}'^{(4)}. \quad (5.25)$$

In Appendix A, we see that $\Delta\tilde{\mathcal{M}}^{(4)}$ can be rewritten such that every term in $\Delta\tilde{\mathcal{M}}'^{(4)}$ merges or clusters all 4 particles into one group and is given in Eq. (A.7). In general, this approach requires defining $\Delta\tilde{\mathcal{M}}'^{(n)}$ such that clustering effects at lower orders ($\Delta\tilde{\mathcal{M}}'^{(k)}$ for $k < n$) are explicitly factored out. However, for $n = 5$ (and higher), this is not the case, as demonstrated in Appendix A. With 5 particles, there is a leftover term that cannot be placed into products of $\Delta\tilde{\mathcal{M}}'^{(k)}$ for $k < 5$, nor does it cluster or merge all 5 particles together. This term must reside in $\Delta\tilde{\mathcal{M}}'^{(5)}$:

$$\Delta\tilde{\mathcal{M}}'^{(5)} \supset 30 c(2)^2 m(\{1, 2\}) [\tilde{w}(2) - \tilde{w}(1)^2] \tilde{w}(1)^3. \quad (5.26)$$

It contains two groups, one of 2 particles and one of 3 particles. Such terms are associated with the map from $\mathcal{P}^c = \{1, 1, 1, 2\} \rightarrow \mathcal{P}^f = \{2, 3\}$, when clustering takes place after the first merging step. For $n \geq 5$, there are terms contributing to $\Delta\tilde{\mathcal{M}}'^{(n)}$ in which each of the n particles cluster or merge with at least one other particle but all n are not grouped together, i.e. $\mathcal{P}^f \neq \{n\}$. We are forced to conclude that Eq. (5.24) does not have any content and is simply a redefinition of Eq. (5.23).

6. Conclusions

In this paper, we developed a formalism to determine the phase space constraints from a clustering algorithm and the effects on a measurement function. This framework can be used to study perturbative corrections that arise from clustering. We expressed the constraints from clustering as a function of the outcome of the algorithm. This is characterized in terms of a partition that specifies which particles are clustered by a clustering function, $c(p)$, and a non-clustering function, $s(\{p\})$, which requires groups of particles not to cluster. We derived a set of unitarity relations on phase space, which can be iteratively applied to eliminate the non-clustering function. As discussed in the text, this requires the introduction of a merging function, $m(\{p\})$, which appears in unitarity relations involving s functions. This rewriting eliminates exclusive phase space constraints that require particles *not* to be in certain regions of phase space, producing only inclusive constraints. This is useful because it simplifies the structure of the most divergent regions of phase space and is a key step in allowing us to relate higher order clustering effects to lower order constraints.

In Sec. 3, we applied this result to the measurement function, weighting the phase space constraints for each outcome of the algorithm by the contribution to the measurement. Grouping terms that have the same phase space constraints from clustering, we rewrote the measurement function purely in terms of inclusive constraints in Eq. (3.23), eliminating the non-clustering function s . This form makes the clustering effects explicit, expressing the measurement function as a term associated with no clustering and a set of corrections. These correction terms are associated with n -particle clustering effects, and are proportional to factors defined in Eq. (3.19) that are the difference between the contribution to $\widetilde{\mathcal{M}}$ with and without clustering. This made the analysis in Sec. 4 of the infrared structure associated with clustering more straightforward. Additionally, we showed that the n particle measurement function $\widetilde{\mathcal{M}}^{(n)}$ can be related to the measurement function appearing at lower orders, $\widetilde{\mathcal{M}}^{(k)}$ with $k < n$, as in Eq. (3.26). This form is particularly useful when trying to understand the higher order structure of jet cross sections.

In Sec. 5 we apply our formalism for the measurement function to examine the perturbative corrections that arise from clustering for a wide class of jet shapes in the soft collinear regime. The main results for the measurement function, chiefly Eq. (3.23) and Eq. (3.26), were used to determine the all-orders structure of the Abelian terms in the soft function, Eq. (5.23). This structure is universal to jet shape observables and depends only on the form of the measurement function and constraints from renormalization. From the all-orders form of the soft function and by count logarithms arising from clustering in the Abelian terms, we found that to all orders clustering NGLs contribute at NLL plus sub-leading logs for the k_T class of algorithms. This same framework can be applied to study non-Abelian terms or non-logarithmic contributions from clustering, since the measurement function contributions from clustering are made explicit.

7. Acknowledgments

We would like to thank Christian Bauer, Christopher Lee, and Matthew Schwartz for help-

ful comments on a draft of this work. This work was supported in part by the Office of High Energy Physics of the U.S. Department of Energy under the Contract DE-AC02-05CH11231 and Grant DE-SC003916. JW was supported in part by a LHC Theory Initiative Postdoctoral Fellowship, under the National Science Foundation grant PHY-0705682.

A. Examples for 4 and 5 Final State Particles

In this appendix we present the measurement function for 4 and 5 final state particles using the formalism of Sec. 2 and Sec. 3. We will start with 4 final state particles, for which there are five possible partitions.

$$\{\{4\}, \{1, 3\}, \{2, 2\}, \{1, 1, 2\}, \{1, 1, 1, 1\}\}. \quad (\text{A.1})$$

As dictated in Eq. (3.23), we split the contribution into two factors: the one associated with \mathcal{P}^c and the one associated with the merging sequence. For each allowed \mathcal{P}^c , the appropriate factors from the first half of Eq. (3.23) are:

$$\begin{aligned} \mathcal{P}^c & : N(\mathcal{P}^c) \prod_{p^c \in \mathcal{P}^c} c(p^c) \Delta \tilde{w}(p^c) \\ \{4\} & : c(4) \Delta \tilde{w}(4) \\ \{1, 3\} & : 4c(3) \tilde{w}(1) \Delta \tilde{w}(3) \\ \{2, 2\} & : 3c(2)^2 \Delta \tilde{w}(2)^2 \\ \{1, 1, 2\} & : 6c(2) \tilde{w}(1)^2 \Delta \tilde{w}(2) \\ \{1, 1, 1, 1\} & : \tilde{w}(1)^4. \end{aligned} \quad (\text{A.2})$$

Each of these factors is then multiplied by $\sum_{\phi \in \mathcal{G}_n(\mathcal{P}^c)} \mathcal{F}(\phi)$. The function $\mathcal{F}(\phi)$, defined in Eq. (3.21), depends on the specific sequence of mergings that is represented by the mapping

$$\phi_c : \mathcal{P}^c = \mathcal{P}_1^\phi \rightarrow \mathcal{P}_2^\phi \cdots \mathcal{P}_j^\phi \rightarrow \mathcal{P}_{j+1}^\phi = \mathcal{P}^f. \quad (\text{A.3})$$

Recall, from the definition of this map, that the first step in this mapping, $\mathcal{P}_1^\phi \rightarrow \mathcal{P}_2^\phi$, must be a merging step and not a clustering clustering. All of the possible mappings for $n = 4$

and the associated factors are listed below.

$\phi_c : \mathcal{P}^c$	\longrightarrow	\mathcal{P}^f	$:$	$\mathcal{F}(\phi)$	
$\{4\}$	\longrightarrow	$\{4\}$	$:$	1	
$\{1, 3\}$	\longrightarrow	$\{4\}$	$:$	$-m(\{1, 3\})$	
$\{2, 2\}$	\longrightarrow	$\{4\}$	$:$	$-m(\{2, 2\})$	
$\{1, 1, 2\}$	\longrightarrow	$\{4\}$	$:$	$-m(\{1, 1, 2\})$	
$\{1, 1, 2\} \longrightarrow \{1, 3\} \longrightarrow$		$\{4\}$	$:$	$2m(\{1, 2\})m(\{1, 3\})$	(A.4)
$\{1, 3\}$	\longrightarrow	$\{1, 3\}$	$:$	1	
$\{1, 1, 2\}$	\longrightarrow	$\{1, 3\}$	$:$	$-2m(\{1, 2\})$	
$\{2, 2\}$	\longrightarrow	$\{2, 2\}$	$:$	1	
$\{1, 1, 2\}$	\longrightarrow	$\{1, 1, 2\}$	$:$	1	
$\{1, 1, 1, 1\}$	\longrightarrow	$\{1, 1, 1, 1\}$	$:$	1	

After combining the factors from Eq. (A.4) with appropriate factor from Eq. (A.2), as dictated by Eq. (3.23), we get

$$\begin{aligned}
\widetilde{\mathcal{M}}^{(4)} = & \widetilde{w}(1)^4 + c(4)\Delta\widetilde{w}(4) + 4c(3)\widetilde{w}(1)\Delta\widetilde{w}(3)[1 - m(\{1, 3\})] \\
& + 3c(2)^2\Delta\widetilde{w}(2)^2[1 - m(\{2, 2\})] + 6c(2)\widetilde{w}(1)^2\Delta\widetilde{w}(2)[1 - m(\{1, 1, 2\})] \\
& - 12c(2)\widetilde{w}(1)^2\Delta\widetilde{w}(2)m(\{1, 2\})[1 - m(\{1, 3\})]
\end{aligned} \tag{A.5}$$

We can write the expression for $\widetilde{\mathcal{M}}^{(4)}$ in the form of Eq. (3.26) by substituting the expressions for $\Delta\widetilde{\mathcal{M}}^{(2)}$ and $\Delta\widetilde{\mathcal{M}}^{(3)}$ given in Eq. (3.27) and Eq. (3.28). The result is

$$\widetilde{\mathcal{M}}^{(4)} = [\widetilde{\mathcal{M}}^{(1)}]^4 + 6\Delta\widetilde{\mathcal{M}}^{(2)}[\widetilde{\mathcal{M}}^{(1)}]^2 + 4\Delta\widetilde{\mathcal{M}}^{(3)}\widetilde{\mathcal{M}}^{(1)} + \Delta\widetilde{\mathcal{M}}^{(4)}, \tag{A.6}$$

where

$$\begin{aligned}
\Delta\widetilde{\mathcal{M}}^{(4)} = & c(4)\Delta\widetilde{w}(4) - 4c(3)\widetilde{w}(1)\Delta\widetilde{w}(3)m(\{1, 3\}) + 3c(2)^2\Delta\widetilde{w}(2)^2[1 - m(\{2, 2\})] \\
& + 6c(2)\widetilde{w}(1)^2\Delta\widetilde{w}(2)[2m(\{1, 2\})m(\{1, 3\}) - m(\{1, 1, 2\})].
\end{aligned} \tag{A.7}$$

Next, we give the primed version as defined under Eq. (5.24):

$$\Delta\widetilde{\mathcal{M}}'^{(4)} = \Delta\widetilde{\mathcal{M}}^{(4)} - 3[\Delta\widetilde{\mathcal{M}}^{(2)}]^2 \tag{A.8}$$

$$\begin{aligned}
= & c(4)\Delta\widetilde{w}(4) - 4c(3)\widetilde{w}(1)\Delta\widetilde{w}(3)m(\{1, 3\}) - 3c(2)^2\Delta\widetilde{w}(2)^2m(\{2, 2\}) \\
& + 6c(2)\widetilde{w}(1)^2\Delta\widetilde{w}(2)[2m(\{1, 2\})m(\{1, 3\}) - m(\{1, 1, 2\})].
\end{aligned} \tag{A.9}$$

Each term is associated with a map with $\mathcal{P}^f \rightarrow \{4\}$.

For $n = 5$, we proceed similarly. There are seven possible partitions, each contributing to the measurement function:

$$\{\{5\}, \{1, 4\}, \{2, 3\}, \{1, 1, 3\}, \{1, 2, 2\}, \{1, 1, 1, 2\}, \{1, 1, 1, 1, 1\}\}. \quad (\text{A.10})$$

The factors associated with clustering are listed below.

$$\begin{aligned} \mathcal{P}^c & : N(\mathcal{P}^c) \prod_{p^c \in \mathcal{P}^c} c(p^c) \Delta \tilde{w}(p^c) \\ \{5\} & : c(5) \Delta \tilde{w}(5) \\ \{1, 4\} & : 5c(4) \tilde{w}(1) \Delta \tilde{w}(4) \\ \{2, 3\} & : 10c(3) c(2) \Delta \tilde{w}(2) \Delta \tilde{w}(3) \\ \{1, 1, 3\} & : 10c(3) \tilde{w}(1)^2 \Delta \tilde{w}(3) \\ \{1, 2, 2\} & : 15c(2)^2 \tilde{w}(1) \Delta \tilde{w}(2)^2 \\ \{1, 1, 1, 2\} & : 10c(2) \tilde{w}(1)^3 \Delta \tilde{w}(2) \\ \{1, 1, 1, 1, 1\} & : \tilde{w}(1)^5 \end{aligned} \quad (\text{A.11})$$

Next we give the factors associated with merging/clustering. For ease of presentation, we give those maps with $\mathcal{P}^f = \{5\}$.

$$\begin{aligned} \phi : \mathcal{P}^c & \longrightarrow \mathcal{P}^f : \mathcal{F}(\phi) \\ \{5\} & \longrightarrow \{5\} : 1 \\ \{1, 4\} & \longrightarrow \{5\} : -m(\{1, 4\}) \\ \{2, 3\} & \longrightarrow \{5\} : -m(\{2, 3\}) \\ \{1, 1, 3\} & \longrightarrow \{5\} : -m(\{1, 1, 3\}) \\ \{1, 1, 3\} \longrightarrow \{1, 4\} \longrightarrow \{5\} & : 2m(\{1, 3\})m(\{1, 4\}) \\ \{1, 2, 2\} & \longrightarrow \{5\} : -m(\{1, 2, 2\}) \\ \{1, 2, 2\} \longrightarrow \{2, 3\} \longrightarrow \{5\} & : 2m(\{1, 2\})m(\{2, 3\}) \\ \{1, 2, 2\} \longrightarrow \{1, 4\} \longrightarrow \{5\} & : m(\{2, 2\})m(\{1, 4\}) \\ \{1, 1, 1, 2\} & \longrightarrow \{5\} : -m(\{1, 1, 1, 2\}) \\ \{1, 1, 1, 2\} \longrightarrow \{1, 4\} \longrightarrow \{5\} & : 3m(\{1, 1, 2\})m(\{1, 4\}) \\ \{1, 1, 1, 2\} \longrightarrow \{1, 1, 3\} \longrightarrow \{5\} & : 3m(\{1, 2\})m(\{1, 1, 3\}) \\ \{1, 1, 1, 2\} \rightarrow \{1, 1, 3\} \rightarrow \{1, 4\} \rightarrow \{5\} & : -6m(\{1, 2\})m(\{1, 3\})m(\{1, 4\}) \\ \{1, 1, 1, 2\} \rightarrow \{1, 1, 3\} \rightarrow \{2, 3\} \rightarrow \{5\} & : -3c(2)m(\{1, 2\})m(\{2, 3\}) \end{aligned} \quad (\text{A.12})$$

The rest of the maps are now given.

$$\begin{array}{llll}
\phi_c : \mathcal{P}^c & \longrightarrow & \mathcal{P}^f & : \quad \mathcal{F}(\phi) \\
\{1, 4\} & \longrightarrow & \{1, 4\} & : \quad 1 \\
\{1, 1, 3\} & \longrightarrow & \{1, 4\} & : \quad -2m(\{1, 3\}) \\
\{1, 2, 2\} & \longrightarrow & \{1, 4\} & : \quad -m(\{2, 2\}) \\
\{1, 1, 1, 2\} & \longrightarrow & \{1, 4\} & : \quad -3m(\{1, 1, 2\}) \\
\{1, 1, 1, 2\} \longrightarrow \{1, 3\} \longrightarrow & \{1, 4\} & : \quad 6m(\{1, 2\})m(\{1, 3\}) \\
\{2, 3\} & \longrightarrow & \{2, 3\} & : \quad 1 \\
\{1, 2, 2\} & \longrightarrow & \{2, 3\} & : \quad -2m(\{1, 2\}) \\
\{1, 1, 1, 2\} \longrightarrow \{1, 1, 3\} \longrightarrow & \{2, 3\} & : \quad 3c(2)m(\{1, 2\}) \\
\{1, 1, 3\} & \longrightarrow & \{1, 1, 3\} & : \quad 1 \\
\{1, 1, 1, 2\} & \longrightarrow & \{1, 1, 3\} & : \quad -3m(\{1, 2\}) \\
\{1, 2, 2\} & \longrightarrow & \{1, 2, 2\} & : \quad 1 \\
\{1, 1, 1, 2\} & \longrightarrow & \{1, 1, 1, 2\} & : \quad 1 \\
\{1, 1, 1, 1, 1\} & \longrightarrow & \{1, 1, 1, 1, 1\} & : \quad 1
\end{array} \tag{A.13}$$

We can construct $\widetilde{\mathcal{M}}^{(5)}$ by combining these factors using Eq. (3.23). After using the definitions of $\Delta\widetilde{\mathcal{M}}^{(2)}$, $\Delta\widetilde{\mathcal{M}}^{(3)}$, and $\Delta\widetilde{\mathcal{M}}^{(4)}$, we get

$$\widetilde{\mathcal{M}}^{(5)} = [\widetilde{\mathcal{M}}^{(1)}]^5 + 10\Delta\widetilde{\mathcal{M}}^{(2)}[\widetilde{\mathcal{M}}^{(1)}]^3 + 10\Delta\widetilde{\mathcal{M}}^{(3)}[\widetilde{\mathcal{M}}^{(1)}]^2 + 5\Delta\widetilde{\mathcal{M}}^{(4)}\widetilde{\mathcal{M}}^{(1)} + \Delta\widetilde{\mathcal{M}}^{(5)}, \tag{A.14}$$

where

$$\begin{aligned}
\Delta\widetilde{\mathcal{M}}^{(5)} = & c(5)\Delta\widetilde{w}(5) - 5c(4)\widetilde{w}(1)\Delta\widetilde{w}(4)m(\{1, 4\}) \\
& + 10c(3)c(2)\Delta\widetilde{w}(2)\Delta\widetilde{w}(3)\left[1 - m(\{2, 3\})\right] \\
& + 10c(3)\widetilde{w}(1)^2\Delta\widetilde{w}(3)\left[2m(\{1, 3\})m(\{1, 4\}) - m(\{1, 1, 3\})\right] \\
& + 15c(2)^2\widetilde{w}(1)\Delta\widetilde{w}(2)^2\left[m(\{1, 4\})m(\{2, 2\}) - 2m(\{1, 2\})(1 - m(\{2, 3\})) - m(\{1, 2, 2\})\right] \\
& + 10c(2)\widetilde{w}(1)^3\Delta\widetilde{w}(2)\left[-3c(2)m(\{1, 2\})m(\{2, 3\}) + 3m(\{1, 4\})m(\{1, 1, 2\})\right. \\
& \quad \left.+ 3m(\{1, 2\})(-2m(\{1, 3\})m(\{1, 4\}) + m(\{1, 1, 3\})) - m(\{1, 1, 1, 2\})\right] \\
& + 30c(2)^2\widetilde{w}(1)^3\Delta\widetilde{w}(2)m(\{1, 2\}).
\end{aligned} \tag{A.15}$$

The primed version is:

$$\begin{aligned}
\Delta\widetilde{\mathcal{M}}'^{(5)} &= \Delta\widetilde{\mathcal{M}}^{(5)} - 10\Delta\widetilde{\mathcal{M}}^{(2)}\Delta\widetilde{\mathcal{M}}^{(3)} \\
&= c(5)\Delta\tilde{w}(5) - 5c(4)\tilde{w}(1)\Delta\tilde{w}(4)m(\{1,4\}) - 10c(3)c(2)\Delta\tilde{w}(2)\Delta\tilde{w}(3)m(\{2,3\}) \\
&\quad + 10c(3)\tilde{w}(1)^2\Delta\tilde{w}(3)\left[2m(\{1,3\})m(\{1,4\}) - m(\{1,1,3\})\right] \\
&\quad + 15c(2)^2\tilde{w}(1)\Delta\tilde{w}(2)^2\left[m(\{1,4\})m(\{2,2\}) + 2m(\{1,2\})m(\{2,3\}) - m(\{1,2,2\})\right] \\
&\quad + 10c(2)\tilde{w}(1)^3\Delta\tilde{w}(2)\left[-3c(2)m(\{1,2\})m(\{2,3\}) + 3m(\{1,4\})m(\{1,1,2\})\right. \\
&\quad \quad \left.+ 3m(\{1,2\})(-2m(\{1,3\})m(\{1,4\}) + m(\{1,1,3\})) - m(\{1,1,1,2\})\right] \\
&\quad + \mathbf{30c(2)^2\tilde{w}(1)^3\Delta\tilde{w}(2)m(\{1,2\})}.
\end{aligned} \tag{A.16}$$

All of the terms in $\Delta\widetilde{\mathcal{M}}'^{(5)}$ are associated with maps that have $\mathcal{P}^f = \{5\}$, except for the last one. This term, whose map is shown in bold in Eq. (A.13), is associated with a map from $\mathcal{P}^c = \{1, 1, 1, 2\}$ to $\mathcal{P}^f = \{2, 3\}$. This contribution is associated with two groups of gluons, one with 2 gluons and the other with 3, rather than from clustering of all 5 gluons. This is the first order at which clustering can occur after a merging step, and it is these terms which cause the difficulty.

References

- [1] A. Abdesselam, E. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth, *et al.*, *Boosted objects: A Probe of beyond the Standard Model physics*, *Eur.Phys.J.C* (2010) [[arXiv:1012.5412](#)].
- [2] A. Altheimer *et al.*, *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, [arXiv:1201.0008](#).
- [3] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, *New clustering algorithm for multi - jet cross-sections in e^+e^- annihilation*, *Phys. Lett.* **B269** (1991) 432–438.
- [4] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant $K(t)$ clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B406** (1993) 187–224.
- [5] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D48** (1993) 3160–3166, [[hep-ph/9305266](#)].
- [6] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [[hep-ph/9707323](#)].
- [7] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [[arXiv:0802.1189](#)].
- [8] **JADE** Collaboration, W. Bartel *et al.*, *Experimental studies on multijet production in e^+e^- annihilation at PETRA energies*, *Z. Phys.* **C33** (1986) 23.
- [9] N. Brown and W. J. Stirling, *Jet cross sections at leading double logarithm in e^+e^- annihilation*, *Phys. Lett.* **B252** (1990) 657–662.

- [10] C. W. Bauer, N. D. Dunn, and A. Hornig, *Subtractions for SCET Soft Functions*, [arXiv:1102.4899](#).
- [11] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *The Soft Function for Exclusive N -Jet Production at Hadron Colliders*, *Phys. Rev.* **D83** (2011) 114030, [[arXiv:1102.4344](#)].
- [12] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion, and J. R. Walsh, *Consistent Factorization of Jet Observables in Exclusive Multijet Cross-Sections*, *Phys. Lett.* **B689** (2010) 82–89, [[arXiv:0912.0262](#)].
- [13] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, *JHEP* **11** (2010) 101, [[arXiv:1001.0014](#)].
- [14] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N -subjettiness*, *JHEP* **03** (2011) 015, [[arXiv:1011.2268](#)].
- [15] M. Jankowiak and A. J. Larkoski, *Jet Substructure Without Trees*, *JHEP* **06** (2011) 057, [[arXiv:1104.1646](#)].
- [16] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N -subjettiness*, [arXiv:1108.2701](#).
- [17] J. R. Walsh and S. Zuberi, *Factorization Constraints on Jet Substructure*, [arXiv:1110.5333](#).
- [18] C. F. Berger, T. Kucs, and G. Sterman, *Event shape / energy flow correlations*, *Phys. Rev.* **D68** (2003) 014012, [[hep-ph/0303051](#)].
- [19] C. W. Bauer, S. Fleming, C. Lee, and G. Sterman, *Factorization of e^+e^- event shape distributions with hadronic final states in Soft Collinear Effective Theory*, *Phys. Rev.* **D78** (2008) 034027, [[arXiv:0801.4569](#)].
- [20] A. Hornig, C. Lee, and G. Ovanessian, *Effective predictions of event shapes: Factorized, resummed, and gapped angularity distributions*, *JHEP* **05** (2009) 122, [[arXiv:0901.3780](#)].
- [21] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *N -Jettiness: An Inclusive Event Shape to Veto Jets*, *Phys.Rev.Lett.* **105** (2010) 092002, [[arXiv:1004.2489](#)].
- [22] S. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tonnesmann, *Jets in hadron-hadron collisions*, *Prog.Part.Nucl.Phys.* **60** (2008) 484–551, [[arXiv:0712.2447](#)].
- [23] G. P. Salam, *Towards jetography*, [arXiv:0906.1833](#).
- [24] A. Banfi and M. Dasgupta, *Problems in resumming interjet energy flows with k_t clustering*, *Phys.Lett.* **B628** (2005) 49–56, [[hep-ph/0508159](#)].
- [25] Y. Delenda, R. Appleby, M. Dasgupta, and A. Banfi, *On QCD resummation with $k(t)$ clustering*, *JHEP* **0612** (2006) 044, [[hep-ph/0610242](#)].
- [26] K. Khelifa-Kerfa, *Non-global logs and clustering impact on jet mass with a jet veto distribution*, [arXiv:1111.2016](#).
- [27] R. Kelley, J. R. Walsh, and S. Zuberi, *Abelian Non-Global Logarithms from Soft Gluon Clustering*, [arXiv:1202.2361](#).
- [28] M. Dasgupta and G. P. Salam, *Resummation of non-global QCD observables*, *Phys. Lett.* **B512** (2001) 323–330, [[hep-ph/0104277](#)].

- [29] M. Dasgupta and G. P. Salam, *Accounting for coherence in interjet $E(t)$ flow: A case study*, *JHEP* **03** (2002) 017, [[hep-ph/0203009](#)].
- [30] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, and S. Marzani, *Non-global logarithms and jet algorithms in high- p_T jet shapes*, *JHEP* **08** (2010) 064, [[arXiv:1004.3483](#)].
- [31] R. Kelley, M. D. Schwartz, R. M. Schabinger, and H. X. Zhu, *Jet mass with a jet veto at two loops and the universality of non-global structure*, [arXiv:1112.3343](#).
- [32] R. Kelley, M. D. Schwartz, and H. X. Zhu, *Resummation of jet mass with and without a jet veto*, [arXiv:1102.0561](#).
- [33] A. Hornig, C. Lee, J. R. Walsh, and S. Zuberi, *Double Non-Global Logarithms In-N-Out of Jets*, [arXiv:1110.0004](#).